NORTHWESTERN UNIVERSITY

I Think I See the Light Curve: The Good (and Bad) of Exoplanetary Inverse Problems

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Physics and Astronomy

By

Joel Colin Schwartz

EVANSTON, ILLINOIS

December 2016

ABSTRACT

I Think I See the Light Curve: The Good (and Bad) of Exoplanetary Inverse Problems

Joel Colin Schwartz

Planets and planetary systems change in brightness as a function of time. These "light curves" can have several features, including transits where a planet blocks some starlight, eclipses where a star obscures a planet's flux, and rotational variations where a planet reflects light differently as it spins. One can measure these brightness changes—which encode radii, temperatures, and more of planets—using current and planned telescopes. But interpreting light curves is an inverse problem: one has to extract astrophysical signals from the effects of imperfect instruments.

In this thesis, I first present a meta study of planetary eclipses taken with the Spitzer Space Telescope. We find that eclipse depth uncertainties may be overly precise, especially those in early *Spitzer* papers. I then offer the first rigorous test of BiLinearly-Interpolated Subpixel Sensitivity (BLISS) mapping, which is widely used to model detector systematics of *Spitzer*. We show that this ad hoc method is not statistically sound, but it performs adequately in many real-life scenarios.

Next, I present the most comprehensive empirical analysis to date on the energy budgets and bulk atmospherics of hot Jupiters. We find that dayside and nightside measurements suggest many hot Jupiters have reflective clouds in the infrared, and that day-night heat transport decreases as these planets are irradiated more. I lastly describe a semi-analytical model for how a planet's surfaces, clouds, and orbital geometry imprint on a light curve. We show that one can strongly constrain a planet's spin axis—and even spin direction—from modest high-precision data. Importantly, these methods will be useful for temperate, terrestrial planets with the launch of the James Webb Space Telescope and beyond.

Acknowledgements

Wow, 5 years it's been now—this still feels *really* surreal...

Let me start by saying that the help I've gotten from Nick, our fellow exocartographers, collaborators, and my other supporters has been...absolutely invaluable, I guess would be the only way to put it. I've said things like that many times now—I'm sure I'll say it again, too. So I want to say stuff to those you won't find elsewhere here.

Friends from Northwestern, McGill, *and* beyond—thanks for being able to share all the hard times, setbacks, long hours, and accomplishments of academia with me. Things don't always work out the way you planned (Go Undercats!), but they do seem to kind of work out in the end. And between the research and the teaching (Therese, Gion, Mark, and Michelle!) and the outreach it's always been really fulfilling.

Family—thanks for always being supportive. There are too many people to list out, but from driving with me to Montreal, to giving me a place to visit and stay, to having parties and holidays at your homes year after year, it's awesome. I love having that to come back to or relax with.

Allison, Carlos, Diana, Kay, Nick, Roger, and Stephanie—thanks for just being great friends. You have no idea how much I love that you guys keep me regular (sounds hilarious read on its own). Maybe we haven't seen each other lately as often as we'd like, but I know we can still talk and laugh about anything when we do. Oh, and Stephanie, the cookies and stuff are *so good* you should be charging!

And Jackie, Cosmin, Jeff, Mom, and Dad—thanks for understanding me in ways that almost no one else can. Whenever we get to spend time together, however that is, I'm always so happy. I know I'm probably all over the place or frustrating sometimes, but you're a big reason why I'm comfortable with myself and whatever I'm doing in the future.

...I wouldn't trade that for anything. $\oint$

Table of Contents

ABSTRACT 3

Acknowledgements 5

List of Tables 10

List of Figures 11

Chapter 1. Introduction and Background 14

1.1. Our (Exo)Planetary Neighborhood 14

1.2. The Wisdom in Light Curves 25

1.3. Tools for Crossing the Void 31

Chapter 2. Features in the Broadband Eclipse Spectra of Exoplanets: Signal or Noise? 39

2.1. Introduction 39

2.2. Broadband Eclipse Spectra 43

2.3. The Significance of Spectral Features 46

2.4. Empirical Estimate of Eclipse Uncertainties 49

2.5. Broadband Spectra with Empirical Uncertainties 56

2.6. Discussion 59

2.7. Conclusions 64

2.8. Interlude I 65

Chapter 3. Knot a Bad Idea: Testing BLISS Mapping for Spitzer Space Telescope Eclipse Observations 66

3.1. Introduction 66

3.2. Optimizing Nuisance Parameters 69

3.3. Synthetic Light Curves 78

3.4. Tests of BLISS 82

3.5. Discussion 109

3.6. Conclusions 114

3.7. Interlude II 116

Chapter 4. Balancing the Energy Budget of Short-Period Giant Planets: Evidence for Reflective Clouds and Optical Absorbers 117

4.1. Introduction 117

4.2. Dayside Energy Budget 120

4.3. Global Energy Budget 128

4.4. Discussion 143

4.5. Updates 153

4.6. Interlude III 160

Chapter 5. Inferring Planetary Obliquity Using Rotational And Orbital Photometry 161

5.1. Introduction 161

5.2. Reflected Light 166

5.3. Kernel Behavior 172

5.4. Discussion 176

5.5. Conclusions 183

References 185

Appendix A. BLISS: Choosing Parameters 209

A.1. Coefficients for the Pixel Sensitivity 209

A.2. Eclipse and Phase Curve 210

Appendix B. Obliquity: Describing Planets and Kernels 211

B.1. Viewing Geometry 211

B.2. Kernel Details 220

List of Tables

2.1 Planets with ≥ 2 thermal eclipse measurements. 43

2.2 Repeated eclipse observations with *Spitzer*. 54

3.1 Works that model *Spitzer* intra-pixel sensitivity with BLISS. 68

4.1 Short-period giant planets with ≥ 2 eclipses at IR wavelengths. 121

4.2 Eclipse depths and phase amplitudes for planets where albedo and heat recirculation can be estimated. 128

4.3 Fitted albedos and heat recirculation efficiencies. 142

4.4 Albedo data for planets with optical eclipses. 148

List of Figures

1.1	A visual overview of exoplanetary science in this thesis.	15
1.2	Mass versus semi-major axis for most confirmed exoplanets.	16
1.3	The Kepler-62 system compared with the Solar System.	18
1.4	Comparing orbital distances in the Solar System with hot Jupiters.	20
1.5	An artist's concept of the hot Jupiter WASP-12b.	22
1.6	Diagrams of methods for detecting exoplanets.	23
1.7	Diagrams of features in the light curves from planetary systems.	27
1.8	The Spitzer, Kepler, and Hubble Space Telescopes.	33
1.9	The future JWST, TESS, and CHEOPS space missions.	36
2.1	Planets as blackbodies: published uncertainties.	47
2.2	Figure 2.1 with empirical uncertainty added in quadrature.	57
2.3	Figure 2.1 with uncertainties inflated by empirical factor.	58
3.1	Marginalization versus optimization for different 2D posteriors.	71
3.2	Example data and residuals from a toy polynomial model.	74
3.3	Example MCMC fits to the data in Figure 3.2.	76
3.4	Example centroid positions and a pixel sensitivity map.	79

3.5	A synthetic light curve and its two components.	83
3.6	Example BLISS knots and how they differs from the true sensitivity.	85
3.7	Eclipse depth posteriors for the light curve in Figure 3.5.	93
3.8	Projected sensitivities for the fits in Figure 3.7.	95
3.9	How BLISS knots vary if optimized or fit directly.	96
3.10	Accuracy/precision of eclipse depths: varied eclipse significances and detector amplitudes.	100
3.11	Figure 3.10 with different amounts of data and BLISS knots.	104
3.12	Figure 3.10 with varied red noise.	106
3.13	Figure 3.10 with different sensitivity maps.	108
4.1	Average emission spectrum for 50 short-period giant planets.	124
4.2	Dayside temperatures for all giant planets with ≥ 2 IR eclipses.	125
4.3	Breakdown of dayside flux for CoRoT-2b at 1.4 μm.	132
4.4	Planetary reflected light for different geometric albedos.	134
4.5	1σ regions in albedo and heat recirculation for thermal planets.	137
4.6	Flux ratios between blackbody and “Sum of Blackbodies” spectra.	139
4.7	Figure 4.5 for all 15 planets in the sample.	141
4.8	Diagram of light curves with and without phase offsets.	154
4.9	Figure 4.5 with and without phase offsets.	157
5.1	A planet’s orbital and rotational apparent albedo.	165

5.2	An example kernel and its defining features.	169
5.3	Contours of kernel properties at first quarter phase.	171
5.4	Longitudinal and latitudinal kernels of a case-study planet.	173
5.5	Kernel width and dominant colatitude throughout an orbit.	175
5.6	Predicted confidence regions for a planet's spin axis.	179
5.7	Apparent albedos of planets with prograde or retrograde spin.	182
B.1	Diagram of a general planetary system with orbital parameters.	212
B.2	Longitudinal and latitudinal kernels of a planet with varied obliquity.	221
B.3	Monte Carlo data on the kernel and apparent albedo.	224

CHAPTER 1

Introduction and Background

To people living in large cities, the night sky probably never looks very impressive—not with the same beauty that astronomy texts can paint, at least. But with clear skies far from light pollution, the view changes completely: thousands of twinkling stars are scattered across constellations and the Milky Way, with hints of many features in the dark. That is a sight to behold, clearly showing the Earth as just one minuscule and precious island in the universe.

It has been easy to consider Earth and our Solar System as special places, and there is still truth in that. This is the only planet we know that harbors life, despite efforts like SETI to seek it elsewhere (Anderson et al., 2002). Yet astonishingly, we now understand that many M-dwarfs and Sun-like stars have planets (Dressing & Charbonneau, 2013; Petigura et al., 2013), often in *multi*-planet systems (e.g. Schneider et al., 2011; Lissauer et al., 2012; Han et al., 2014). This is the age of exoplanetary science (Figure 1.1), and our Solar System is far less alone than ever imagined. That these new worlds have been hiding in (relatively) plain sight, essentially forever, shows just how difficult it can be to find them!

1.1. Our (Exo)Planetary Neighborhood

Latham et al. (1989) discovered and Cochran et al. (1991) verified the first *candidate* planetary system, HD 114762, but the latter authors could not rigorously determine the

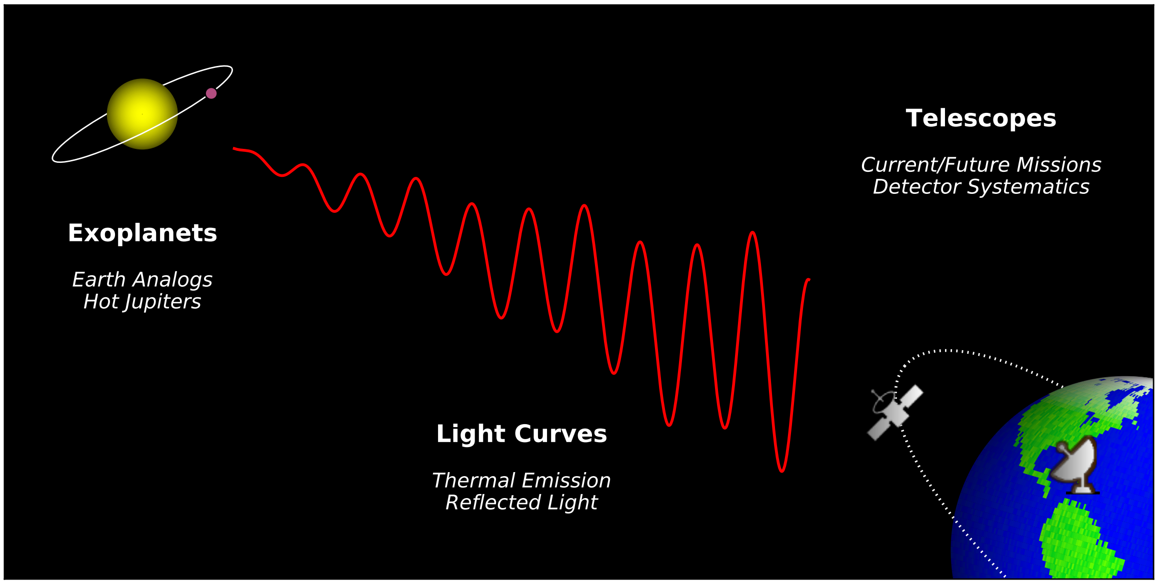


Figure 1.1. A visual overview: exoplanets and methods of detecting them (Section 1.1), light curves and their features (Section 1.2), and the telescopes used for observing (Section 1.3). Satellite icons modified from Freepik on Flaticon.com.

companion object's mass. The first *confirmed* discoveries of exoplanets came in 1994 (PSR B1257+12 system; Wolszczan et al., 1994) and 1995 (51 Pegasi b; Mayor & Queloz, 1995), and the catalog of nearby planets has grown considerably since. In fact, the nearest confirmed planet is only about 4.2 ly away (Proxima Centauri b; Anglada-Escudé et al., 2016). While many other worlds are on the order of 10^3 ly from Earth, they are still relatively close given the expanse of our galaxy ($\sim 10^5$ ly across).

There are now thousands of confirmed planets in our galactic neighborhood (Schneider et al., 2011) and even more candidates (e.g. Batalha et al., 2013; Mullally et al., 2015). As Figure 1.2 shows, these planets span the gamut of our Solar System and more: gas giants with large semi-major axes (e.g. HR 8799 System; Marois et al., 2008; Marley et al., 2012), potential ice giants around the size of Neptune (e.g. Kepler-421b; Borucki et al.,

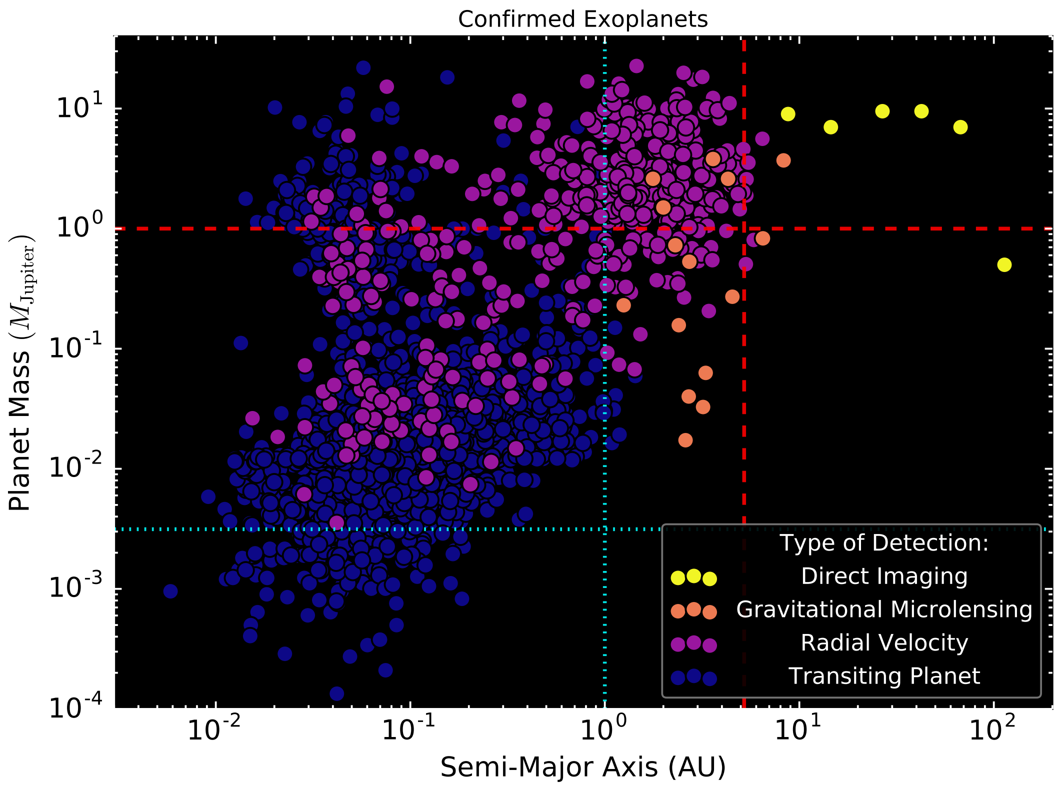


Figure 1.2. The mass and semi-major axis of most confirmed exoplanets ($\sim$ 2900), plotted on a log–log scale. The colors show how each object was detected (see below in Section 1.1): yellow for direct imaging, orange for gravitational microlensing, purple for radial velocity, and indigo for transiting planets. The dashed red and dotted cyan lines show the values for Jupiter and Earth, respectively. Parameters taken from the Exoplanets Data Explorer on Exoplanets.org (Han et al., 2014).

2011; Kipping et al., 2014), and terrestrial worlds similar in size to our inner planets (e.g. Kepler-37b and -37c; Batalha et al., 2013; Barclay et al., 2013).

Some exoplanets have *no* known analogs in the Solar System. Consider "super-Earths" (e.g. GJ 1214b; Charbonneau et al., 2009): these are worlds with masses between that of Earth and Uranus (i.e. $M_\oplus \leq M \leq 14.5\ M_\oplus$; they are also called "mini-Neptunes" near the more massive end). Models of planet formation did not exclude super-Earths from existing (e.g. Papaloizou & Terquem, 2005), but discovering such planets was still surprising. And with evidence that orbits of distant Kuiper Belt objects may be influenced by a body of at least 10 $M_\oplus$ (Batygin & Brown, 2016), the so-called "Planet 9" of our Solar System *could* wind up being a mini-Neptune!

Early studies into super-Earths looked at their internal structure (Valencia et al., 2006), the mass-radius relationships of both GJ 876d (Valencia et al., 2007a) and solid exoplanets (Seager et al., 2007), and inferring their bulk properties (Valencia et al., 2007b). These works showed that, for a given mass, there is a maximum radius a *rocky* super-Earth can have—bigger planets have a large amount of water or a significant H/He envelope. More recently, Rogers (2015) modeled that when radii are about 1.62 $R_\oplus$ or greater, *at least* half of planets this size are not dense enough to be made purely of iron and silicates. As also shown by Weiss & Marcy (2014), a planet's mass can vary considerably for a given radius, so super-Earths can have many different compositions.

Super-Earths (and mini-Neptunes) are more common than larger gas giants (e.g. Dong & Zhu, 2013; Han et al., 2014), orbit in multi-planet systems roughly 40% of the time (Rowe et al., 2014), and are often ordered by radius when a planet the size of Neptune or larger is present in the system (Ciardi et al., 2013). Rocky planets with higher masses will also have stronger surface gravities and so retain atmospheres better. Ginzburg et al. (2015) found an ideal region in mass and temperature where planets accrete atmospheres but do not become gas giants—many observed super-Earths are in this category. As with exoplanets in general, most super-Earths are discovered around Sun-like (e.g. Kepler-20; Fressin et al., 2011; Borucki et al., 2011; Gautier III et al., 2012) and near-Sun-like stars (e.g. Kepler-62; Borucki et al., 2011, 2013). In fact, both Kepler-62e and -62f are *within*

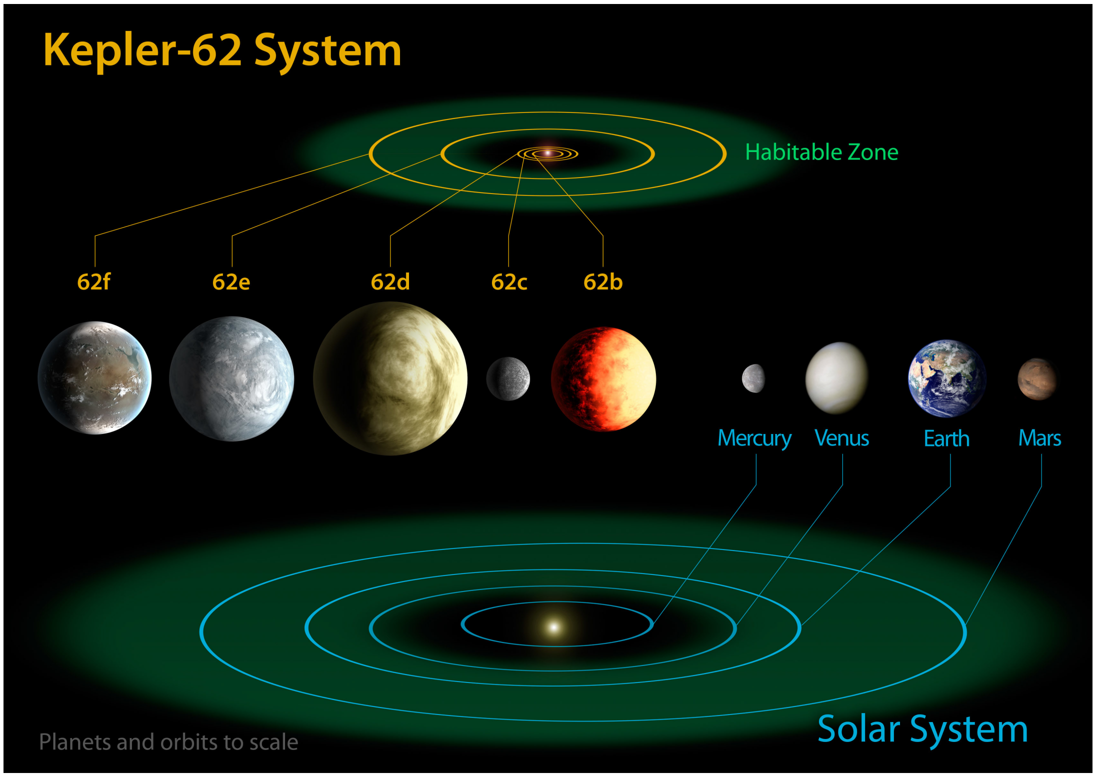


Figure 1.3. The planets orbiting Kepler-62 compared with the inner planets of our Solar System. All orbital distances and planetary radii are to scale; the Kepler-62 planets are artists' concepts. Note that Kepler-62e and -62f are inside the habitable zone (green band), where planets with enough atmospheric pressure could have liquid water on their surfaces. Image credit to NASA.

the habitable zone of their host star (Figure 1.3), or orbital distances where liquid water could exist on a planet's surface.

Since most main-sequence stars are cooler than the Sun (e.g. LeDrew, 2001) and burn at least an order of magnitude longer (Tinsley, 1980), M-dwarfs could be ideal stars to search for habitable planets around. Gillon et al. (2016) recently found three Earth-sized planets orbiting the ultra-cool dwarf TRAPPIST-1, two of which are near the inner edge of the habitable zone. Better yet, Anglada-Escudé et al. (2016) discovered a super-Earth in the habitable zone around the red dwarf Proxima Centauri—the nearest star to the Sun—though its habitability is debated (e.g. Barnes et al., 2016; Martin et al., 2016; Meadows et al., 2016; Ribas et al., 2016). Whether or not M-dwarfs can support a planet with life is a complex question (cf. Barnes et al., 2013; Yang et al., 2013), but the sheer numbers alone are exciting for someday characterizing an Earth-like planet.

However, observing small, rocky planets in detail is difficult. The contrast ratio for an Earth-like planet around a Sun-like star is $\sim 10^{-10}$ at visible wavelengths (Bailey, 2014). This is beyond the reach of current telescopes, but research is progressing towards that threshold (e.g. Trauger & Traub, 2007; Cheng-Chao et al., 2015). For now, we are better suited to characterize bright targets we *can* detect—often the exotic "hot Jupiters."

The Solar System has 4 known giant planets at large semi-major axes, where they can accrete gas envelopes from the protoplanetary disk. As shown in Figure 1.4, though, hot Jupiters are found extremely close to their host stars (e.g. WASP-8b; Queloz et al., 2010), sometimes over 25$\times$ closer than Mercury's orbit (e.g. WASP-43b; Hellier et al., 2011)! During planet formation the inner nebula should be too hot for volatiles to condense (e.g. Chambers, 2004), and disk instability should cause giant planets to form at large semi-major axes anyway (e.g. Boss, 2000). Thus, hot Jupiters likely have short-period orbits through either disk migration (e.g. Lin et al., 1996; Papaloizou et al., 2007) or high-eccentricity migration (Rasio & Ford, 1996; Fabrycky & Tremaine, 2007). But in any case, hot Jupiters are usually bright compared with their host stars in the optical and infrared (e.g. HD 189733b; Bouchy et al., 2005; Deming et al., 2006; Evans et al., 2013), are extreme examples of a physical laboratory (Heng & Showman, 2014), and orbit $\sim 1\%$ of nearby Sun-like stars (Wright et al., 2012).

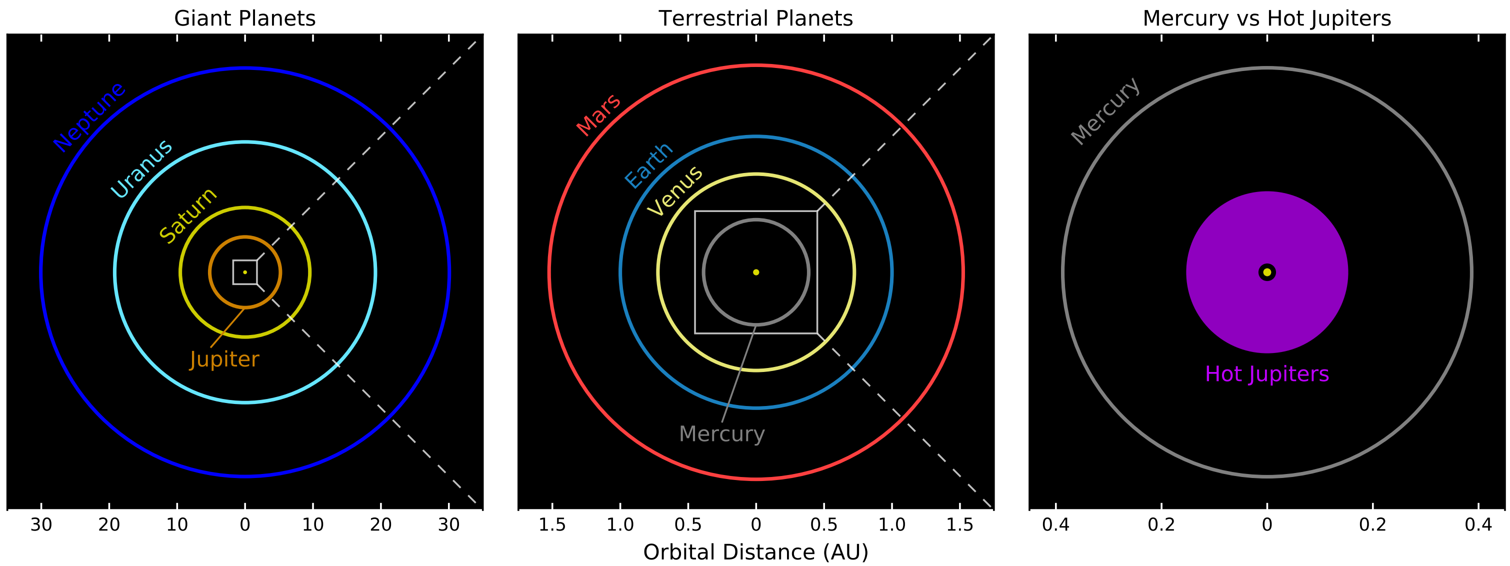


Figure 1.4. The average orbital distances for planets in the Solar System (left and center panels) compared to most hot Jupiters (purple region in right panel). Note that the panels zoom in from left to right (gray boxes and dashed lines). The colors of the Solar System orbits roughly match the actual planets. The Sun is the yellow circle in each panel (size not to scale). Hot Jupiters orbit *very* close to their host stars.

When a massive planet has a short-period orbit, the tidal forces—which raise and lower oceans on the Earth—become very strong. These forces are so extreme that hot Jupiters should be tidally locked to their host stars on timescales of $\sim 10^6$ years (i.e. very quickly; Goldreich & Soter, 1966). Just as we see only one side of the Moon from Earth, hot Jupiters have permanent day and nightsides, where energy from the star only reaches the nightside by transport through the atmosphere (e.g. Perez-Becker & Showman, 2013). Unsurprisingly, these daysides can be upwards of ~ 3000 K (e.g. Cowan & Agol, 2011b) before even considering sources of internal heat! In fact, some hot Jupiters start losing atmospheric mass due to tidal dissipation (e.g. Figure 1.5, of WASP-12b; Li et al., 2010) or because they are irradiated so much (cf. Baraffe et al., 2004; Hubbard et al., 2007). And, these planets' magnetic fields may not protect their envelopes from early, intense stellar winds (Grießmeier et al., 2004).

Exoplanets span multiple orders of magnitude in radius, mass, and semi-major axis. How would one find them all? Figure 1.6 summarizes the primary methods used to find the planets in Figure 1.2 (i.e. the colors). For planets bright enough and widely separated from their host stars, direct imaging (far left panels) is a natural choice (e.g. Beta Pictoris b; Lagrange et al., 2009; Chauvin et al., 2012). Here one distinguishes a planet's light from the star using a starshade or coronagraph. This lets one observe the planet at a (potentially) large fraction of orbital phases, no matter how its orbital plane is seen from Earth. However, direct imaging works best for big planets with large semi-major axes—it is not practical yet for hot Jupiters or colder terrestrial worlds (Bowler, 2016).

Gravitational microlensing (center left panels of Figure 1.6) is an alternative that does not rely on light *from* the planet. Instead, the brightness of a background star increases

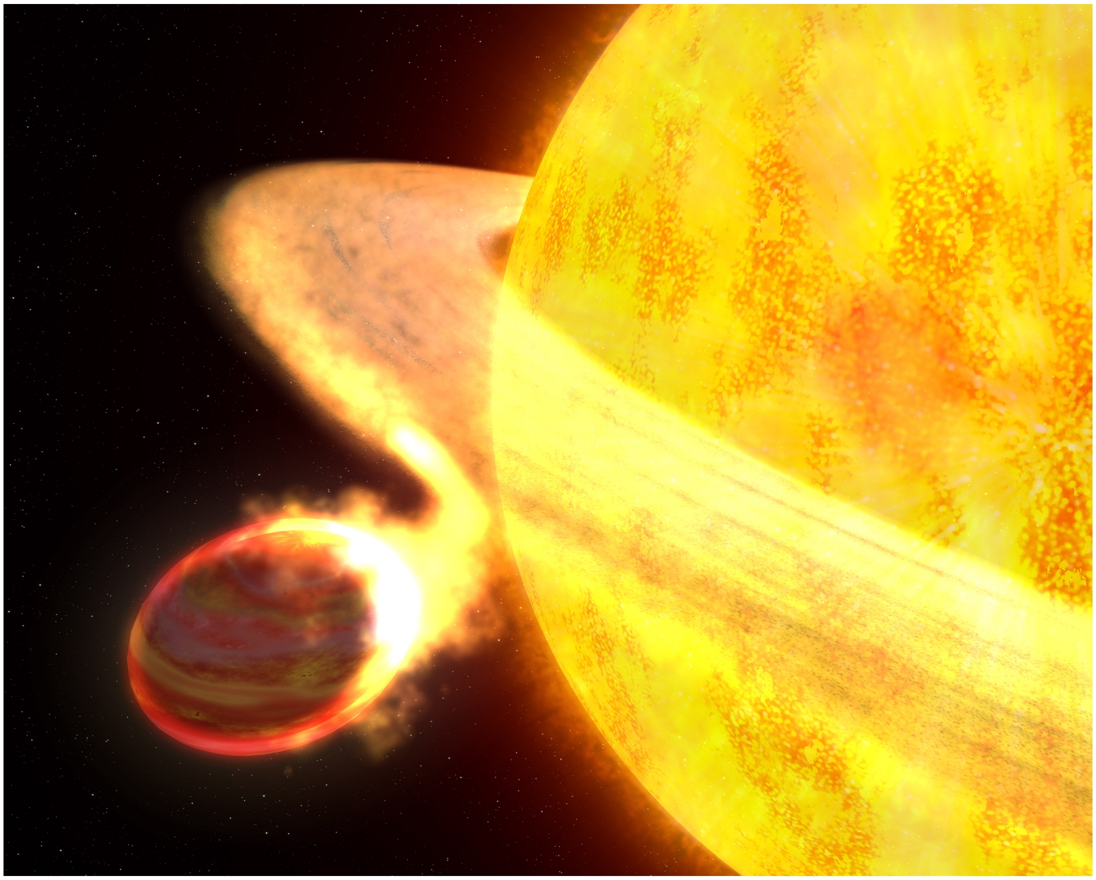


Figure 1.5. An artist's concept of the hot Jupiter WASP-12b. With a semi-major axis only $\sim 3.1\times$ its host star's radius (Chan et al., 2011), WASP-12b has a prolate (i.e. egg-like) shape. Tidal dissipation is stripping mass from its atmosphere, meaning there is likely some disk of planetary gas surrounding the star (Li et al., 2010). Image credits to NASA, ESA, and G. Bacon.

when a planetary system passes in front of it (e.g. Beaulieu et al., 2006). But, this method has only uncovered a small number of planets (so far; orange points in Figure 1.2).

For massive planets, one can measure the radial velocity (center right panels of Figure 1.6) of the host star (Wright & Gaudi, 2013). Stars and planets orbit their common center of mass, so light from the star will get more red- and blue-shifted the more edge-on we view the planet's orbital plane from Earth. This method can find planets ill-suited for

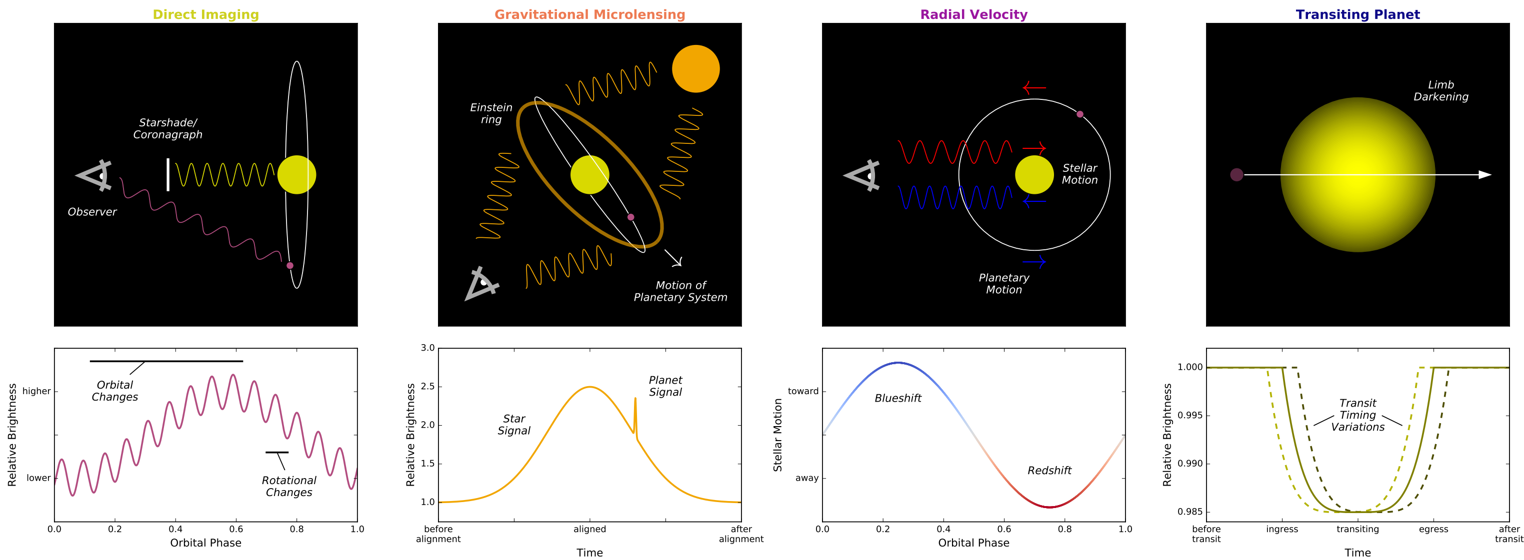


Figure 1.6. Diagrams of methods for detecting exoplanets (upper row) and examples of the signals seen (lower row). Each method's title is the same color as planets found by that method in Figure 1.2. *Far Left:* In direct imaging, light from the star is blocked out, leaving only flux from the planet. Here one can see both orbital and rotational changes in the planet's brightness (Section 1.2). *Center Left:* In gravitational microlensing, a planetary system passes in front of a background star (orange), bending the light and increasing its brightness. The planet can alter this signal a small amount more than expected by the star alone. *Center Right:* With radial velocity, a planet causes a star to wobble, red- and blue-shifting the stellar light over the planet's orbit. More massive planets can affect the same star's motion more, depending on how the planet's orbital plane is aligned to Earth. *Far Right:* For transiting planets, some stellar light is blocked when the planet passes in front of the star. In multi-planet systems, transits can happen earlier or later than expected (transit timing variations or TTVs) because of how the planets gravitationally tug on one another.

direct imaging, but that are massive enough to tug significantly on their host stars (e.g. ~ 75 m/s for 55 Cancri b; Butler et al., 1997). Until the mid-2000s, this was the most common way to discover exoplanets (Schneider et al., 2011).

If a planet's orbit is very edge-on, it will repeatedly pass in front of its star (i.e. transit, like Mercury or Venus seen from Earth) and block some of the light (e.g. OGLE-TR-56b; Konacki et al., 2003). This transit method (far right panels of Figure 1.6) is now the typical way planets are found (indigo points in Figure 1.2; Han et al., 2014). As with any detection method, transits have biases: it is easiest to discover large planets with small semi-major axes—like hot Jupiters—and short-period planets are generally the fastest to confirm because the time between transits is short. For comparison, extraterrestrials could only confirm seeing Earth with transits after staring at the Sun for *at least* one of our years.

There is another possibility for multi-planet systems: just as planets can make their stars wobble, they can also tug on each other and change their orbital speeds. This causes transit timing variations (TTVs; e.g. Miralda-Escudé, 2002; Holman & Murray, 2005; Agol et al., 2005), where a given planet crosses its star slightly earlier or later than expected (far lower right panel of Figure 1.6). Patterns in TTVs can help one learn about other planets in a system, even those that do *not* transit(e.g. KOI-872; Nesvornỳ et al., 2012).

In short, exoplanets are ubiquitous in the Milky Way and (almost surely) other galaxies, many are nothing like the planets (or moons) of our Solar System, and there are several ways to discover them. The connecting thread is the light we see from these stars and their planets, and indeed "light curves" are a basis of this dissertation. How we interpret these observations can both improve and hinder our understanding of exoplanets—so next we look at light curves in more detail.

1.2. The Wisdom in Light Curves

Simply put, a light curve is the brightness of an object over time, be that a distant galaxy or supernova or pulsar. Stars and their orbiting planets are particularly challenging to observe: the contrast between these bodies is large, and planets can have inhomogeneous surface and atmospheric features. There are two main ways planets contribute to stellar light curves.

Everything—animals included—radiates electromagnetic (EM) energy at wavelengths that depend mostly on the object's temperature (i.e. as a blackbody). The light from very irradiated planets is dominated by this thermal emission. Energy from the host star is absorbed by the atmosphere and/or surface then re-radiated at longer wavelengths (Seager & Deming, 2010), though the greenhouse effect can block some of this outgoing radiation (e.g. Venus; Pollack et al., 1980). On cooler planets, a source of internal energy could add to the thermal emission (e.g. for Juptier; Hanel et al., 1981). Planets have equilibrium temperatures between ~ 3000 K down to ~ 50 K, so one should see them radiate thermally in the infrared, or wavelengths longer than $\sim 1\ \mu$m (e.g. Seager & Deming, 2010; Cowan & Agol, 2011b; Bailey, 2014). This thermal light can vary from isotropic (i.e. same in all directions) to just dayside emission (i.e. due to tidal locking; some hot Jupiters), and hotter planets are easier to detect this way.

On the other hand, a planet's atmosphere or surface can be reflective. This reflected light from the host star will dominate the planet's flux at wavelengths where thermal emission is negligible. For G-, K-, and M-type stars of roughly 6000–3000 K, their peak EM radiation is between about 0.5–1.0 μm, or in the red-optical (Tinsley, 1980). Radiation is scattered off atmospheric molecules or surface features of the planet, meaning reflected light is best seen at optical (and possibly shorter) wavelengths (e.g. Seager & Deming, 2010; Heng & Demory, 2013).

The line between thermal and reflected light gets blurred for very hot planets. This can happen because the planet's thermal emission leaks into the optical (e.g. Heng & Showman, 2014; Bailey, 2014), or the planet is fairly reflective in the near-infrared (e.g. hot Jupiters; Schwartz & Cowan, 2015). In contrast, the outer planets and moons of our Solar System are very cold and thus only *reflect* light in the visible.

Thus, distinguishing the flux from some planets often means tackling an inverse problem. A forward problem usually has one unique solution, such as summing a planet's thermal emission and reflected light (i.e. $A + B = C$ where $\{A, B\}$ are known). Inverse problems are typically less constrained, like when *inferring* a planet's thermal and reflected light from only its total flux (i.e. $C = A + B$ where C is known). Recognizing these inverse problems is part of the story. This thesis focuses on finding and interpreting some answers to inverse problems spurred by the light curves of planetary systems.

Simplistically, then, planets emit thermally at long wavelengths (i.e. infrared) and reflect stellar light at short wavelengths (i.e. optical). Both mechanisms affect the light curve of the planetary system in several ways, as we show in Figure 1.7. The upper panel shows a transiting planet on the solid orbit and a world that does not transit on the

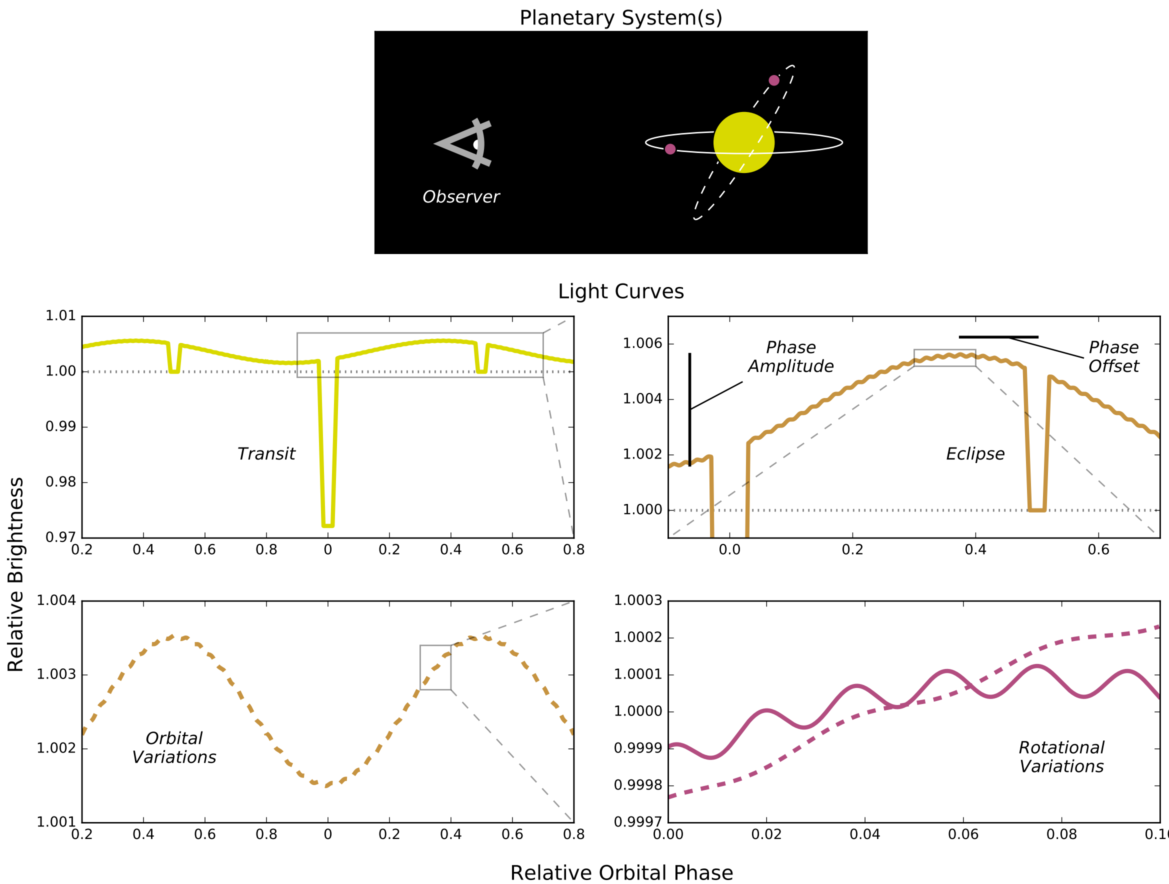


Figure 1.7. Diagrams of the light curves from planetary systems at an arbitrary wavelength. *Upper Panel:* An observer and example planets on circular orbits around a star. The transiting planet follows the solid curve while the one that does not transit is on the dashed curve. *Other Panels:* Example light curves that these bodies create at different levels of detail (yellow to orange to purple curves)—note how the panels zoom in (gray boxes and dashed lines). The curve styles match the orbits in the upper panel, and the flux from the star alone is shown as the dotted line. Key features in each panel are indicated. Both curves and axes in the lower right panel have been shifted to help comparison. One can learn about a planet and its atmosphere by analyzing the changes in brightness of that planetary system.

dashed orbit. These line styles are used in the other panels—we refer to these diagrams throughout the rest of this section.

For transiting planets (middle left panel), the biggest way the planet affects the light curve is the transit itself. One infers the planet's orbital inclination (i.e. tilt of its orbital plane seen from Earth) from the transit duration, the planet's radius from the transit depth (i.e. difference in stellar brightness before and during transit), and the orbital period and so semi-major axis from the time between transits. The shape of the transit (i.e. roundness of bottom) also constrains the star's limb darkening, or how stars look dimmer near the edge of their disk. Above all, if the transit depth varies at different wavelengths, the planet likely has an atmosphere blocking more starlight in some bands (Charbonneau et al., 2006; Haswell, 2010). By studying light passing through the atmosphere during a transit, one can learn which molecules are in the planet's envelope (e.g. Bean et al., 2013; Stevenson et al., 2014b).

A transiting planet also passes *behind* its host star half an orbit later (assuming zero eccentricity)—this is an eclipse. As shown in the middle right panel, these dips in brightness are shallower than transits because only the planet's light is blocked (e.g. Seager & Deming, 2010). Eclipses help rule out false positives of planets, and by comparing the eclipse and transit depths, one learns the planet's intensity relative to the star at a given wavelength. This can be converted to a brightness temperature at eclipse, useful for describing the dayside in the infrared (e.g. Cowan & Agol, 2011b). At visible wavelengths, an eclipse depth measures a planet's geometric albedo, or how reflective it is at full phase (e.g. Seager & Deming, 2010; Heng & Demory, 2013). This is good for constraining scattering (e.g. Sudarsky et al., 2000) or energy absorption (e.g. Angerhausen et al., 2015). One can even map hotspots by observing the planet as it moves into and out of eclipse (e.g. Majeau et al., 2012).

The phase amplitude is a finer feature (middle right panel). The dayside gradually comes into view as the planet orbits from transit to eclipse—vice versa for the nightside (Cowan et al., 2015). This means the brightness of the star will appear to rise and fall throughout the planet's orbit ("phase variations"). One can combine the phase amplitude with the eclipse and transit depths to estimate a planet's nightside temperature at given infrared wavelengths (e.g. Cowan & Agol, 2011b; Stevenson et al., 2014c).

On planets with an atmosphere, winds move the absorbed energy around, and can change where the hottest region is. For tidally locked planets, one sometimes finds that the peak brightness of the light curve occurs at a time other than eclipse (middle right panel). This is called a phase offset (Heng & Showman, 2014), and seeing it means the planet's nightside is brighter at that wavelength than one might otherwise expect.

The thermal features described above let us predict bulk properties of transiting planet atmospheres: how much total radiation they absorb and how efficiently that energy is transported (this is an inverse problem; Cowan & Agol, 2011b; Schwartz & Cowan, 2015). These properties can then help one infer the molecules an atmosphere is composed of (e.g. Sudarsky et al., 2000; Heng & Demory, 2013) or the wind patterns (e.g. Showman et al., 2010; Heng et al., 2011).

What about planets—particularly terrestrials—that are not in edge-on orbits? Directly imaging a terrestrial planet around a Sun-like star is difficult (Bailey, 2014), but consider the reflected light from such a rocky body as it orbits (lower left panel). The light curve will show orbital variations (Oakley & Cash, 2009), similar to the phase variations above for transiting planets. In this case light scatters from different latitudes over the planetary year, helping one map a planet from North to South and figure out its orbital geometry (Kawahara & Fujii, 2010, 2011; Fujii & Kawahara, 2012). Seasonal cycles can also make the planet's reflection change over time (Robinson et al., 2010), just as some regions of Earth look different in Summer versus Winter. For specular (i.e. glint-like) reflection, orbital light curves could help one detect a planet's oceans or icecaps (cf. Robinson et al., 2014; Cowan et al., 2012a).

Now think about directly imaging a spinning planet for just a few *days* (lower right panel). Here the light curve is influenced finely by the planet's rotational variations (Ford et al., 2001). How this planet looks—where continents, oceans, ices, and maybe clouds are—affects its brightness at any moment because given structures reflect light differently at each wavelength. That means rotational light curves can help one infer how long the planet's day is (Pallé et al., 2008), its colors (e.g. Fujii et al., 2010; Cowan & Strait, 2013), and its longitudinal map (Cowan et al., 2009). If the brightness changes day to day, it could also mean the planet has variable clouds or weather (Pallé et al., 2008; Schwartz et al., 2016). Naturally, constraining a planet's map and orbital geometry from its yearly and daily light curves is also an inverse problem.

This all ties back to finding planets "hospitable" for life, beyond merely considering the habitable zone. Indeed, a planet could be hospitable by having a thicker atmosphere (Seager, 2013), different types and amounts of elements in its atmosphere (Goldblatt, 2016), or even having a large axial tilt (Williams & Kasting, 1997). This dissertation is a small step towards describing habitability in comprehensive terms. Although hot Jupiters are a big focus because they are (relatively) easy to observe, our methods will also apply to terrestrial worlds (e.g. Koll & Abbot, 2015). And this benefits finding future targets with the best properties for supporting life—we now look at bridging the colossal space between Earth and her distant cousins.

1.3. Tools for Crossing the Void

Techniques in planetary science are influenced by which planets are interesting. Studying the Earth is relatively easy: we *live* on and take samples from this planet, and routinely fly satellites around it to monitor the atmosphere, continents, and oceans (e.g. Earth Observing System; Winker et al., 2003; Xiong et al., 2009; Roy et al., 2014). Reaching other parts of the Solar System is trickier, yet we have sent spacecraft like *Cassini-Huygens* to Saturn (Matson et al., 2003) and Titan (Lebreton et al., 2005), *MESSENGER* to Mercury (Leary et al., 2007), and the Voyager probes to our ice giants (Stone & Miner, 1986, 1989) and beyond (e.g. Borovikov & Pogorelov, 2014). We have even visited and brought back samples from our Moon (e.g. Team, 1969), a momentous achievement. But considering the distance to exoplanets—and that no one knew what Pluto looked like until last year (*New Horizons*; e.g. Stern et al., 2015)—remote sensing becomes extremely vital. Here a single *pixel* is worth thousands of words!

Again, exoplanets and their host stars are usually observed in the visible (i.e. reflected light) and infrared (i.e. thermal emission). Contending with Earth's greenhouse gases and atmospheric distortion (e.g. Jacob, 1999) means that most data comes from space-based telescopes, especially in the infrared. The bulk of these measurements are photometric, where all photons from an object in a band of wavelengths are counted simultaneously (e.g. Deming et al., 2007; Todorov et al., 2009; Baskin et al., 2013). The alternative is spectroscopy, where light is split into many wavelength channels (e.g. Tinetti et al., 2010; Schwarz et al., 2015). This reveals planetary atmospheres in more detail, showing features like high-altitude hazes (e.g. HD 189733b; Sing et al., 2011) or water abundances (e.g. WASP-43b; Kreidberg et al., 2014). Spectroscopy can be turned into photometry by summing the flux from all the wavelength channels, but photometric data on exoplanets tends to be more common (e.g. Bailey, 2014).

For space missions, placing a telescope near a Sun-Earth Lagrangian point (e.g. Koon et al., 2008) or in an Earth-trailing orbit (e.g. Van Dyk et al., 2013) lets one observe some planets continuously. Other instruments are limited by their low Earth orbits if in space or Earth's rotation if on the ground. Viewing a planet's phase variations is easier with continuous data (e.g. Cowan et al., 2012b), whereas multiple eclipses or transits are often observed separately but modeled in parallel (e.g. Agol et al., 2010; Deming et al., 2015). Data on single eclipses may not be as robust, particularly in the early days of observing with a given instrument (cf. Hansen et al., 2014; Ingalls et al., 2016).

Space-based missions have been valuable to this dissertation and exoplanet science in general. Prime among these is the Spitzer Space Telescope (upper left of Figure 1.8), launched by NASA in 2003 (Werner et al., 2004). As part of the Great Observatories program, *Spitzer* was not designed to characterize planets, but still covers important bands for thermal emission. Its InfraRed Array Camera (IRAC; Fazio et al., 2004) has four channels—3.6, 4.5, 5.8, and 8.0 μm—though only the shortest two work since the "warm" phase began in 2009 when its helium coolant ran out (Dunbar, 2009). *Spitzer* has taken a large majority of the eclipse and phase data we use (e.g. Beerer et al., 2010; Todorov et al., 2012; Wong et al., 2015).

Figure 1.8. Significant space telescopes in the study of exoplanets: *Spitzer* at upper left, *Kepler* at upper right, and *Hubble* at the bottom. Image credits to NASA.

At visible wavelengths the Kepler Space Telescope (upper right of Figure 1.8) is key, operating since 2009. *Kepler* is basically a large light bucket, designed to stare at a patch of the galactic plane (i.e. near Cygnus and Lyra) and collect photons between roughly 0.42–0.9 μm. The benefit is that the brightness of more than 10^5 stars could be observed simultaneously over a span of several years (Koch et al., 2010). The *Kepler* team has confirmed over 2300 exoplanets, including 1284 in the latest release (Morton et al., 2016). While 2 of the spacecraft's 4 reaction wheels have been damaged, *Kepler* now uses radiation pressure from the Sun to stabilize its field of view along the ecliptic for ~ 2.5 months at a time (K2 mission; Howell et al., 2014). Our best statistics on planets orbiting Sun-like stars (e.g. Batalha et al., 2013; Silburt et al., 2015) come from *Kepler* transit light curves. For this thesis, eclipses observed with *Kepler* have helped constrain the albedos of hot Jupiters.

The Hubble Space Telescope (bottom of Figure 1.8) is going strong for transit spectroscopy of planets, and Deming et al. (2016) advocates using *Hubble* for more eclipse and phase curve observations, too. Its Wide Field Camera 3 (WFC3; Kimble et al., 2008) typically takes planetary measurements from ~ 1.1–1.7 μm but can observe down to ~ 0.2 μm. As with *Spitzer*, *Hubble* was not originally intended to observe exoplanets, yet could keep operating through the mid-2030s until being de-orbited or boosted (Wall, 2015). And, numerous other missions have either contributed to studying exoplanets (e.g. CoRoT; Baglin et al., 2007) or soon will (e.g. *Gaia*; Lindegren et al., 2007).

There are several future missions that will push planetary science forward, as well. The flagship is the James Webb Space Telescope (JWST; Gardner et al., 2006), on schedule to launch in October 2018 (left of Figure 1.9). The spacecraft will orbit the Sun-Earth

L2 point, where a five-layer sunshield will passively cool it to below 50 K (Gardner et al., 2006). Like *Hubble* or *Spitzer*, JWST is a multi-purpose observatory that partly will characterize exoplanets—headed by NASA with help from the European and Canadian Space Agencies. As *Spitzer* and *Hubble* probe the thermal structure of hot Jupiters, JWST should advance this technique (even to a few temperate super-Earths) with its multiple high-resolution, infrared instruments (e.g. NIRSpec, MIRI; Beichman et al., 2014).

Meanwhile, the Transiting Exoplanet Survey Satellite (TESS; Ricker et al., 2015) is slated to launch ahead of JWST in late 2017 (center of Figure 1.9). TESS will pave the way by taking light curves inside 0.6–1.0 μm of more than 2×10^5 dwarf stars (types F5–M5), much brighter than those seen by *Kepler*. This all-sky survey is scheduled to last two years, and should find hundreds of planets smaller than Neptune that JWST can follow-up on (Ricker et al., 2015). Also set for 2017 by the ESA is the CHaracterizing ExOPlanets Satellite (CHEOPS; Broeg et al., 2013), seen at the right of Figure 1.9. CHEOPS will do photometry between 0.4–1.1 μm of $\sim$ 500 bright targets where high-precision radial velocity is possible—it can even detect Earth-sized transits.

Further out (2020–30s), a larger flagship mission would image planets in even finer detail. The proposals include a Large UltraViolet-Optical-InfraRed Surveyor (LUVOIR; Kouveliotou et al., 2014), an Advanced Technology Large-Aperture Space Telescope (ATLAST; Postman et al., 2010), or a High-Definition Space Telescope (HDST; Dalcanton et al., 2015). Such missions would use 8–16 m diameter mirrors and have angular resolutions of $\sim$ 0.01 arcseconds, or 5–10$\times$ better than *Hubble*! But while *Hubble* could resolve Earth from the Sun at a distance of $\sim$ 10–20 parsecs, it cannot image at the contrast level between these bodies (e.g. Bailey, 2014). Instead, the concept telescopes could probe the

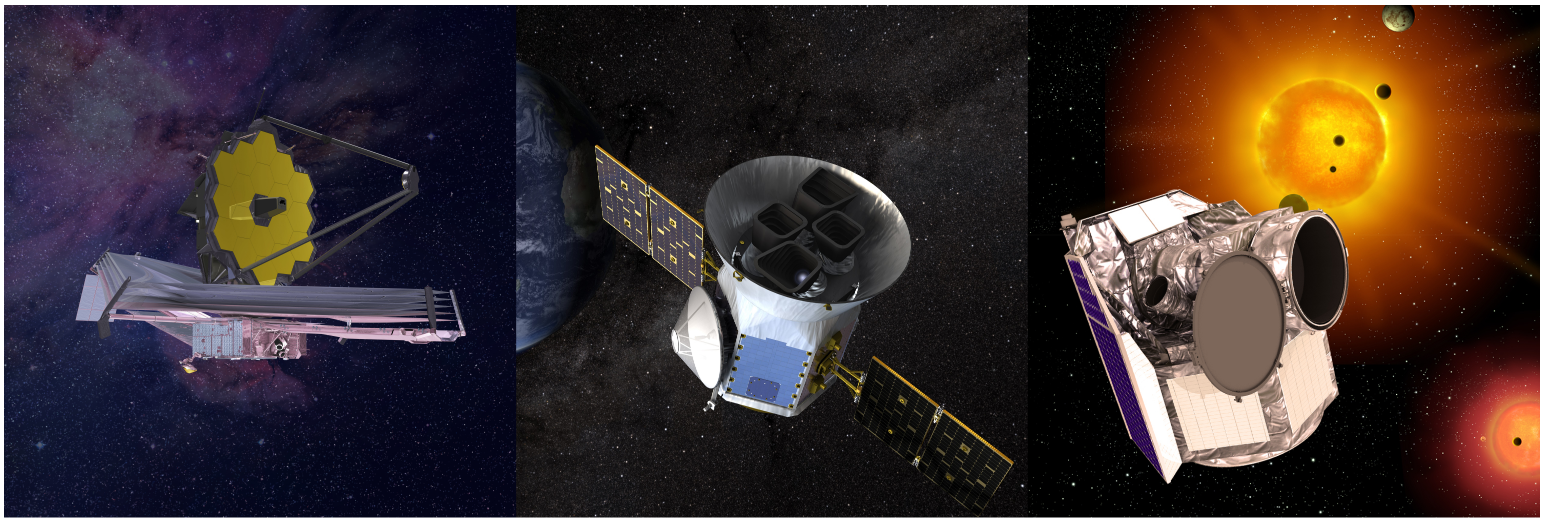


Figure 1.9. Artist concepts of some future missions for exoplanet observing: the James Webb Space Telescope (JWST) at left, the Transiting Exoplanet Survey Satellite (TESS) at center, and the CHaracterizing ExOPlanets Satellite (CHEOPS) at right. Image credits to NASA for JWST/TESS, and ESA for CHEOPS.

atmosphere of an Earth-twin at high contrast using either an internal coronagraph or independent starshade.

These mature, fledgling, and unhatched missions keep planetary science relevant well into the future. But despite anyone's efforts, telescopes run into troubles during their lives. Some are unforeseen, such as the primary mirror on *Hubble* focusing wrongly and needing corrective optics (Burrows et al., 1991) or broken reaction wheels on *Kepler* that cost the spacecraft its original targets (Howell et al., 2014). Others are known in advance: a battery heating cycle on *Spitzer* makes the optics expand and contract, moving the target on the detector in a "sawtooth" pattern (Grillmair et al., 2012). In fact, unfolding JWST on its way to the Sun-Earth L2 point is complicated (Gardner et al., 2006), and it will be crucial to avoid (major) mistakes because no maintenance can be done after launch. None of this is unique to exoplanets of course—even the cosmic microwave background was shrouded at first by *pigeons* nesting inside a radio horn antenna in New Jersey (e.g. Singh, 2010)!

When telescopes do work as designed, their detectors or cameras can have issues instead. Photometry generally uses many charge-coupled devices (CCDs; e.g. *Kepler*) to record the light from objects. That flux can be hard to quantify accurately for planetary systems because their images are often very pixelated (e.g. Stevenson et al., 2012a). Each CCD can get saturated with charge if exposed too long (e.g. Barbe, 1975) and so underestimate flux from a target, though software typically prevents this for space telescopes. Pixels may also exhibit a ramp, meaning they become more sensitive over time (e.g. Deming et al., 2006).

Worse, the sensitivity *within* one pixel can vary and plague targets that move on the detector (leading to another inverse problem; e.g. Ballard et al., 2010; Crossfield et al., 2012b), or specific CCDs can die and must be excluded from analyses. Even with healthy pixels, absolute photometry at the level of planetary eclipses is tricky (e.g. $\sim 0.1\%$, versus $\sim 1\%$ stability for *Spitzer*; cf. Bailey, 2014; Reach et al., 2005). This means the relative flux within one light curve may be fine, but cannot be directly compared to another.

What follows, then, is a variety of research into light curves, the planets (and stars) that make them, and intricacies in teasing signals from noise. Physics is often divided into theory and experiment, but this dissertation breaks that barrier and lies both between and throughout. These are not definitive answers to all planetary inverse problems—that is hilarious at best—but they do add proverbial stones to the ever-expanding castle of planetary knowledge.

One can picture this thesis as stepping through a possible chronology of exoplanets. Chapter 2 shows that claims of interesting physics and chemistry in planetary atmospheres may be premature—I created the first figures for this study on past eclipse depths. Chapter 3 shifts to the present, where we numerically test synthetic light curves with a popular method (called BLISS mapping) to model intra-pixel sensitivities in *Spitzer* data. In Chapter 4, we compile eclipse and phase data on hot Jupiters to better characterize the current empirical trends in these planets' bulk atmospherics. We conclude in Chapter 5 with a semi-analytical look at why light curves encode planetary maps and spin properties—and how we may utilize these facts with future observations.

CHAPTER 2

Features in the Broadband Eclipse Spectra of Exoplanets: Signal or Noise?

Note: I contributed the first versions of the figures to this study, as well as helped with the literature review to gather thermal eclipse data.

This chapter is adapted from Hansen, C. J., Schwartz, J. C., and Cowan, N. B. 2014, MNRAS, 444, 3632.

2.1. Introduction

An exoplanet on an edge-on orbit periodically passes behind its host star. The decrement in thermal flux that occurs during such an eclipse is a measure of the dayside brightness temperature of the planet. The brightness temperature of a planet varies with wavelength, primarily because of the atmosphere's wavelength-dependent opacity and vertical temperature profile (e.g., Deming et al., 2005; Seager et al., 2005; Barman et al., 2005; Burrows et al., 2007, 2008; Fortney et al., 2008; Knutson et al., 2008; Désert et al., 2009). If different wavelengths probe the same atmospheric layer (e.g., a cloud deck) then the planet will appear to have a blackbody spectrum. In the absence of clouds, a planet may still have a blackbody spectrum if the atmospheric layers probed are isothermal. Indeed, the emission spectra of some planets are reported to be featureless: e.g., TrES-2 (O'Donovan et al., 2010), TrES-3 (Fressin et al., 2010), WASP-18b (Nymeyer et al., 2011), and WASP-12b (Crossfield et al., 2012a).

In principle, the detection of molecular bands in the infrared emission spectrum of a planet enables the retrieval of greenhouse gas abundances and the vertical temperature profile of the planet (e.g., Madhusudhan & Seager, 2009; Madhusudhan et al., 2011; Lee et al., 2012; Line et al., 2012). Spectral resolution is critical to such retrieval exercises because a high-resolution emission spectrum is more likely to deviate significantly from a blackbody, and renders the retrieval problem well-constrained. This bodes well for current and future efforts to perform *bona fide* emission spectroscopy. So far, however, the vast majority of exoplanet emission measurements have been broadband eclipse photometry.

A typical retrieval model uses a dozen parameters to describe the atmospheric composition and vertical temperature profile, while a typical hot Jupiter has only been observed in 2–4 thermal broadbands. Even for the few planets with 6 or 7 thermal measurements, the photometric retrieval problem is under-constrained.

A widely noted consequence of the parameter–data mismatch is that exact atmospheric properties cannot be uniquely determined, making color-color and color-magnitude diagrams more realistic approaches to atmospheric classification (Baskin et al., 2013; Beatty et al., 2014; Triaud, 2014).

The less-discussed aspects of under-constrained retrieval are that (1) there is no way to reject erroneous measurements, and (2) the estimated uncertainties on eclipse depths directly affect the uncertainties on atmospheric parameters. This is in stark contrast to over-constrained problems such as fitting an occultation model to time-series data, for which it is customary to perform outlier rejection (e.g., σ-clipping), and for which the photometric uncertainties are estimated in the process of fitting a model to the data, rather than trusting the output of aperture photometry routines.

Nonetheless, many exoplanet discoveries have been based on broadband emission spectra: a temperature inversion in the atmosphere of HD 209458b was inferred from 4 broadband eclipse depths (Knutson et al., 2008), disequilibrium chemistry was invoked to explain the 6-band emission spectrum of GJ 436b (Stevenson et al., 2010), and high atmospheric C/O was discovered based on 7 broadband eclipses of WASP-12b (Madhusudhan et al., 2011). These successes have led to classifying planets based on temperature inversions (using 2 broadbands per planet; Knutson et al., 2010) and C/O ratio (using ≥ 4 bands per planet; Madhusudhan, 2012).

Temperature inversions and non-solar chemistry have since been disputed for each of the exemplar planets due to the re-reduction of existing data (Beaulieu et al., 2011) acquisition of new data at the same wavelength (Cowan et al., 2012b; Zellem et al., 2014) or acquisition of new data at different wavelengths (Crossfield et al., 2012a). Such challenges are not unique to eclipse *photometry*: the featureless day-side emission spectrum of HD 189733b (Grillmair et al., 2007) exhibited an absorption feature at a later epoch (Grillmair et al., 2008),[1] and line emission from the dayside of HD 189733b (Swain et al., 2010) has been disputed by Mandell et al. (2011).

Nor are issues of repeatability limited to superior conjunction: the first thermal phase measurements of an exoplanet (Harrington et al., 2006) were later found to be off by 80° in phase and more than a factor of 2 in amplitude (Crossfield et al., 2010); the first half of the thermal phase measurements of Knutson et al. (2007b) were later found to be corrupted by detector systematics (Agol et al., 2010).

[1]Although this was interpreted as evidence of planetary variability, that hypothesis is inconsistent with the more extensive monitoring campaign of Agol et al. (2010).

The situation is similar for transit spectroscopy, where initial claims of molecular absorption bands (Tinetti et al., 2007; Swain et al., 2008; Tinetti et al., 2010) were disputed on the basis of data reduction, error estimation, and astrophysical variability (Ehrenreich et al., 2007; Désert et al., 2009; Gibson et al., 2011; Désert et al., 2011c; Crouzet et al., 2012).

Indeed, Burrows (2014) offers a sobering review of the exoplanet atmospheric characterization field, speculating that many of the extraordinary claims of the past decade may be overturned by better data.

In this article we attempt to reconcile Burrows' pessimistic view with the growing body of papers making statements about planetary atmospheres based on a handful of eclipse measurements. Instead of focusing on a single planet, we perform a holistic analysis of all transiting planets with multiple eclipse measurements. We consider only broadband measurements (for which it is easy to quantify the number of independent observational constraints) of eclipse depths (which are unaffected by star spots). Our approach is to compare the goodness-of-fit and evidence for three classes of models: blackbodies, self-consistent radiative transfer, and spectral retrieval. Since the disputes over atmospheric properties have often revolved around the reliability of eclipse depths, we empirically estimate the accuracy of broadband eclipse measurements. Notably, the dominant "signal" in space-based eclipse photometry is usually the detector sensitivity, which must be modeled using the very same observations of the science target.

Future observations of transiting planets with the James Webb Space Telescope are likely to resolve many of the current scientific disputes about the nature of hot Jupiter atmospheres. Attempts to push the observatory to smaller and cooler planets, however, will still rely on self-calibration; error estimation and repeatability will therefore remain critical issues.

2.2. Broadband Eclipse Spectra

A search on exoplanet.org (Wright et al., 2011) combined with a careful literature review yields 44 exoplanets with published photometric eclipse measurements in at least two thermal wavelengths ($\lambda > 1$ μm), summarized in Table 2.1. In most cases, only a single occultation has been measured at each wavelength. Bolded numbers signify measurements based on more data: multiple eclipses and/or an eclipse embedded in phase variations.

Table 2.1. Planets with at least 2 thermal eclipse measurements.

Planet	Wavelengths (μm)
CoRoT-1b	1.65, 2.15, 3.6, 4.5
CoRoT-2b	2.15, 3.6, 4.5, 8.0
GJ 436b	3.6, 4.5, 5.8, **8.0**, 16.0, 24.0
HAT-P-1b	3.6, 4.5, 5.8, 8.0
HAT-P-2b	3.6, 4.5, 5.8, 8.0
HAT-P-3b	3.6, 4.5
HAT-P-4b	3.6, 4.5
HAT-P-6b	3.6, 4.5
HAT-P-7b	3.6, 4.5, 5.8, 8.0
HAT-P-8b	3.6, 4.5
HAT-P-12b	3.6, 4.5
HAT-P-23b	2.15, 3.6, 4.5
HD 149026b	3.6, 4.5, 5.8, **8.0**
HD 189733b	2.15, **3.6**, **4.5**, 5.8, **8.0**, 16.0, 24.0
HD 209458b	2.15, 3.6, **4.5**, 5.8, 8.0, **24.0**
KELT-1b	3.6, 4.5
Kepler 5b	3.6, 4.5
Kepler-6b	3.6, 4.5
Kepler-12b	3.6, 4.5
Kepler-13Ab	2.15, 3.6, 4.5

Planet	Wavelengths (μm)
Kepler-17b	3.6, 4.5
TrES-1b	3.6, 4.5, 8.0
TrES-2b	2.15, 3.6, 4.5, 5.8, 8.0
TrES-3b	1.25, 2.15, 3.6, 4.5, 5.8, 8.0
TrES-4b	3.6, 4.5, 5.8, 8.0
WASP-1b	3.6, 4.5, 5.8, 8.0
WASP-2b	3.6, 4.5, 5.8, 8.0
WASP-3b	3.6, 4.5, 8.0
WASP-4b	2.15, 3.6, 4.5
WASP-5b	**1.25**, **1.65**, **2.15**, 3.6, 4.5
WASP-8b	3.6, 4.5, 8.0
WASP-12b	1.25, 1.65, 2.15, **3.6**, **4.5**, 5.8, 8.0
WASP-14b	3.6, 4.5
WASP-17b	4.5, 8.0
WASP-18b	**3.6**, **4.5**, 5.8, 8.0
WASP-19b	1.65, 3.6, 4.5, 5.8, 8.0
WASP-24b	3.6, 4.5
WASP-33b	2.15, 3.6, 4.5
WASP-43b	3.6, 4.5
WASP-48b	1.65, 2.15, 3.6, 4.5
XO-1b	3.6, 4.5, 5.8, 8.0
XO-2b	3.6, 4.5, 5.8, 8.0
XO-3b	3.6, **4.5**, 5.8, 8.0
XO-4b	3.6, 4.5

The data in Table 2.1 are taken from the following references: (Agol et al., 2010; Alonso et al., 2010; Anderson et al., 2010, 2011; Barnes et al., 2007; Baskin et al., 2013; Beatty et al., 2014; Beaulieu et al., 2011; Beerer et al., 2010; Blecic et al., 2013; Campo et al., 2011; Charbonneau et al., 2005, 2008; Chen et al., 2014; Christiansen et al., 2010; Cowan et al., 2012b; Croll et al., 2010a,b,c; Crossfield et al., 2012a; Cubillos et al., 2013; Deming et al., 2005, 2006, 2007, 2011; Demory et al., 2007; Désert et al., 2011b,a; Fortney et al., 2011; Fressin et al., 2010; Gillon, M. et al., 2009; Gillon et al., 2010; Knutson et al., 2007a,b, 2008, 2009c,a,b, 2012; Lewis et al., 2013; López-Morales et al., 2010; Machalek et al., 2008, 2009, 2010; de Mooij et al., 2013; Nymeyer et al., 2011; O'Donovan et al., 2010; O'Rourke et al., 2014; Richardson et al., 2003; Rogers et al., 2009; Rostron et al., 2014; Shporer et al., 2014; Smith et al., 2012; Stevenson et al., 2010, 2012a, 2014a; Todorov et al., 2009, 2012, 2013; Wheatley et al., 2010; Wong et al., 2014; Zellem et al., 2014; Zhou et al., 2013).

Since we are merely concerned with the emergent spectra of the bodies at superior conjunction, it is immaterial if a planet has an eccentric orbit (GJ 436b, HAT-P-2b, WASP-8b, WASP-14b, XO-3b) or is a highly-irradiated brown dwarf (KELT-1b). The majority of these observations—in particular, all those longward of 3 μm— were made with the Spitzer Space Telescope (Werner et al., 2004). In cases where multiple values have been published, we adopt the most recent.

We fit a blackbody spectrum to the eclipse depths for each planet using the published transit depth and stellar effective temperature. We assume symmetric, Gaussian, error bars for the eclipse depths; in the few cases were asymmetric error bars were published, we take the mean of the upper and lower error bars. The transit depth and stellar effective temperature have associated uncertainties that tend to have a gray impact on the planet's spectrum and hence we neglect them in the current analysis.

In the interest of simplicity, we ignore the detector spectral response functions and instead compute the Plank function at the central wavelength of each photometric observation. Moreover, by using the stellar effective temperature rather than a detailed stellar model, we are treating the star as a blackbody. These assumptions are reasonable for broadband measurements in the infrared.

2.3. The Significance of Spectral Features

A spectral retrieval model can provide a better fit to observations than a blackbody, because it has roughly a dozen free parameters, rather than one. A self-consistent radiative transfer model lies somewhere in between, with a few variables. In order to compare the evidence for these models, we use the Bayesian Information Criterion (BIC; Schwarz et al., 1978). BIC is a simple way to compare the evidence for models with different numbers of parameters: $BIC = \chi^2 + k \ln N$, where χ^2 is the usual badness-of-fit, k is the number of free parameters and N is the number of data. It is similar in spirit to the reduced χ^2 in that it penalizes models with many parameters, but it remains well-defined when there are fewer data than there are parameters, as is the case for current photometric eclipse retrieval. The Akaike Information Criterion (AIC; Akaike, 1974) penalizes complex models even more than the BIC for $N < 7.4$, i.e., for all of the planets considered here. Moreover, Chen & Chen (2008) note that both BIC and AIC tend to be biased in favor of complex models in the small-N, large-k regime. In short, our use of the BIC gives models with many free parameters the benefit of the doubt.

As a baseline, we fit a blackbody and compute the BIC for each planet in our sample using the published eclipse depths and uncertainties. The only unknown is the blackbody temperature, so $k = 1$ and $BIC_{\rm BB} = \chi^2_{\rm BB} + \ln N$. Figure 2.1 shows the blackbody BIC plotted against the number of wavebands available for each planet. Gray denotes the quality of a blackbody fit: the dashed gray line is a perfect fit to a blackbody ($\chi^2_{\rm BB} = 0$), while the gray region denotes a good fit ($\chi^2_{\rm BB}/N \approx 1$ with 68.3% confidence interval).

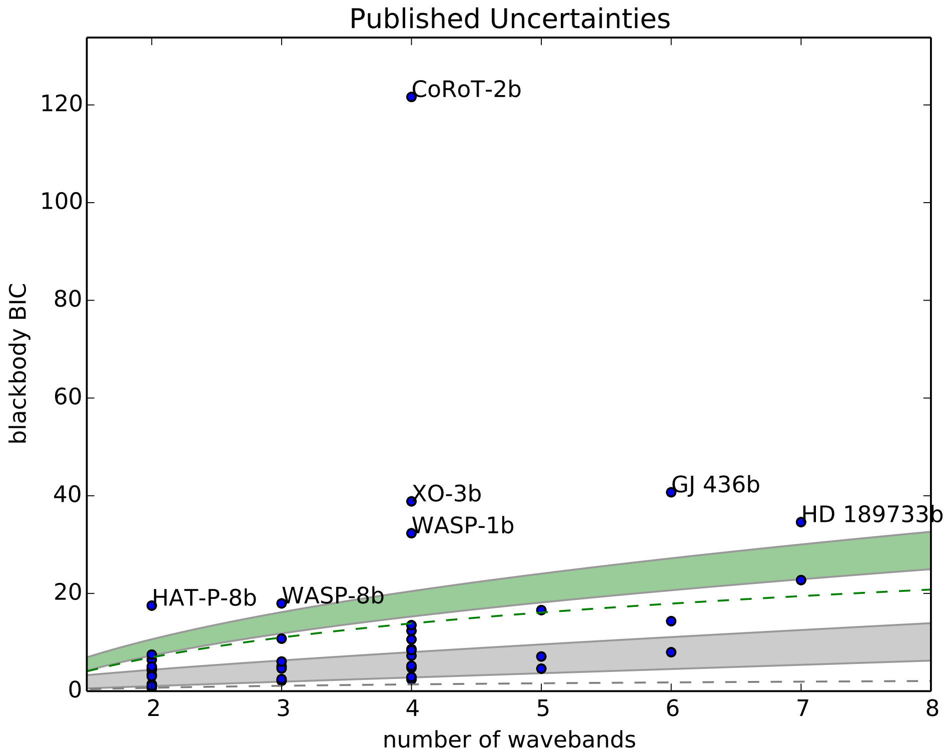


Figure 2.1. The Bayesian Information Criterion (BIC) of a blackbody fit is plotted against the number of thermal wavebands for which photometric eclipse measurements have been obtained; each dot represents one of the 44 transiting planets in our sample. The dashed gray line is a perfect fit to a blackbody ($\chi^2_{\rm BB} = 0$), while the gray region denotes a good fit ($\chi^2_{\rm BB}/N \approx 1$ with 68.3% confidence interval). Planets that lie well above the gray region are poorly fit by a blackbody; the vertical distance above the gray indicates the strength of broadband features in that planet's emission spectrum. Green denotes the quality of a hypothetical spectral retrieval fit: the dashed line is a perfect fit ($\chi^2_{\rm SR} = 0$), while the green region is a good fit ($\chi^2_{\rm SR}/N \approx 1$ with 68.3% confidence interval). Planets that lie in or above the green region may favor spectral retrieval, if published uncertainties are taken at face value.

Since there are few data, the χ^2 distribution is broad and asymmetrical, with a tail towards large values (the colored swaths denote the 68.3% intervals of the χ^2 distribution). Planets that lie well above the gray region are poorly fit by a blackbody, given the published uncertainties. The vertical distance above the gray indicates the strength of broadband features in that planet's emission spectrum. CoRoT-2b exhibits by far the most featured broadband emission spectrum of any transiting planet, a fact not lost on observers (e.g., Cowan et al., 2011).

We also consider an idealized spectral retrieval model with 10 free parameters: 6 parameters for the vertical temperature–pressure profile and 4 for molecular abundances (Madhusudhan & Seager, 2009). Some recent retrieval studies have 2 additional abundance variables, for a total of 12 model parameters (e.g., Stevenson et al., 2014a), so our adoption of 10 is conservative. Since it is under-constrained, one might expect spectral retrieval to provide perfect fits to broadband emission spectra (i.e., $\chi^2_{\rm SR} = 0$). We denote this scenario with the dashed green line in Figure 2.1 ($BIC_{\rm SR} = 10 \ln N$). In practice, spectral retrieval involves *a priori* constrains (e.g., priors on plausible chemistry) so their fits have been in the range $\chi^2_{\rm SR}/N = 0.5$–2 (Madhusudhan & Seager, 2010; Madhusudhan et al., 2011; Madhusudhan, 2012). We therefore also plot a green region denoting $BIC_{\rm SR} = \chi^2_{\rm SR} + 10 \ln N$ (i.e., a spectral retrieval fit with $k = 10$ and $\chi^2_{\rm SR}/N \approx 1$).

We expect that spectral retrieval would produce BIC values in the green swath. While the derivation of BIC relies on assumptions that may not be entirely valid for spectral retrieval, planets that lie above the green region exhibit a preference for spectral retrieval as compared to a blackbody fit ($BIC_{\rm SR} < BIC_{\rm BB}$). For example, CoRoT-2b has been well fit using spectral retrieval ($\chi^2_{\rm SR}/N = 0.725$; Madhusudhan, 2012); if the published eclipses are taken at face value, then there is very strong evidence that spectral retrieval is a better model than a blackbody for this planet.

If published eclipse values and uncertainties are taken at face value, then many hot Jupiters lie above the gray region, indicating that they are poorly fit by blackbodies, but below the green region, implying that the poorly-fitting blackbody is favored over spectral retrieval, according to the BIC. While one could perform spectral retrieval on these data and conceivably obtain interesting atmospheric constraints, they should be taken with a grain of salt because spectral retrieval is probably the *wrong model* given the current data.

Figure 2.1 shows seven planets with broadband emission spectra that invite a full spectral retrieval: CoRoT-2b, GJ 436b, HAT-P-8b, HD 189733b, WASP-1b, WASP-8b, and XO-3b. This list includes a few of the best/brightest transiting targets in GJ 436b, HD 189733b, and XO-3b. Since the Poisson (photon-counting) noise is smaller for bright targets, the smaller error bars might reveal intrinsic molecular bands present in planetary emission spectra. Alternatively, the eclipse uncertainties for bright targets may be dominated by systematic error rather than Poisson noise. Since it is notoriously difficult to estimate systematic errors (Topping & Worrell, 1957), it is critical to empirically evaluate the eclipse accuracy via repeated measurements (Lyons, 1992).

2.4. Empirical Estimate of Eclipse Uncertainties

The instruments currently used to measure exoplanet eclipses are pushed orders of magnitude beyond their design specifications for the simple reason that transiting short-period planets were not known to exist when the instruments were designed (e.g., the 2% stability of *Spitzer* IRAC; Fazio et al., 2004). The raw photometry therefore suffers from detector systematics that are comparable to, and sometimes dwarf, the astrophysical signal of interest (e.g., Charbonneau et al., 2005; Deming et al., 2005).

In what follows, we focus on *Spitzer* because a) 133 of 154 published broadband thermal eclipse measurements were obtained with this telescope, b) these observations have the smallest quoted uncertainties and hence place the strongest constraints on atmospheric structure and composition, and c) these are essentially the only thermal eclipse measurements to have been repeated.

New observing modes with *Spitzer* have improved the data quality over the past decade: staring rather than dithering, only observing in a single waveband at a time, increasing the frequency of the heater cycling, and the peak-up method for keeping the target centroid on the same region of a pixel throughout long observations. Furthermore, there have been improvements in our understanding of *Spitzer* systematics, especially for large data-sets, including pixel-by-pixel ramp correction (Knutson et al., 2007a), polynomial decorrelation (Knutson et al., 2008), double-exponential ramp correction (Agol et al., 2010), Gaussian decorrelation (Ballard et al., 2010), BLISS mapping (Stevenson et al., 2012a), and use of the noise pixel (Knutson et al., 2012; Lewis et al., 2013). It is now routine for combined detector$\times$astrophysics models to fit the data within 10–20% of Poisson noise.

Despite excellent fits, residuals usually exhibit red (time-correlated) noise. The wavelet-based method of Carter & Winn (2009) has been used to estimate the impact of red noise on eclipse depth uncertainties (e.g., the full-orbit phase curves of HD 189733b; Knutson et al., 2012), and Independent Component Analysis (Waldmann, 2012) has been used to perform blind signal de-mixing for transit spectroscopy (Waldmann et al., 2013). Although these methods are better motivated than quick-and-dirty methods such as residual binning and residual permutation (Cowan et al., 2012b), none seem to produce accurate error bars in numerical tests: uncertainty estimates are still too small in the presence of red noise and naïve methods often perform best (Cubillos et al., 2014).

In order to avoid these subtleties of error estimation we would like to fit the data so well that there is no red noise in the residuals. This drives observers to use increasingly complex models. It is notable that the current leading detector models for *Spitzer* channels 1 & 2 are non-parametric (Ballard et al., 2010; Knutson et al., 2012; Stevenson et al., 2012a; Lewis et al., 2013). This is commonly taken to mean that they have *no* free parameters, but it might be more accurate to say that they have a large, but vague, number of parameters.[2] One therefore has to be wary of over-fitting, and should strive to compare models of varying complexity in a Gaussian framework.

Instead of debating the merits of detector models and uncertainty estimation schemes, we now consider the empirical accuracy of eclipse measurements.

2.4.1. Parallel Analysis of Multiple Eclipses

The ideal way to determine the uncertainty on a measurement is to repeat it: obtain many (> 2) eclipse measurements and their standard deviation should be a robust measure of the eclipse uncertainty. This exercise has been performed five times with *Spitzer*: 6 eclipses of HD 189733b at 8 μm (Agol et al., 2010), 11 eclipses of GJ 436b

[2]The Gaussian decorrelation scheme of Knutson et al. (2012) and Lewis et al. (2013) has an effective number of detector parameters roughly equal to the area of the centroid range, $\Delta x \Delta y$ divided by the Gaussian smoothing area, $\sigma_x \sigma_y$. This quantity is typically in the hundreds.

at 8 μm (Knutson et al., 2011), 4 eclipses of 55 Cancri e at 4.5 μm (Demory et al., 2012), 3 eclipses of HD 209458b at 24 μm (Crossfield et al., 2012b), and 12 eclipses of XO-3b at 4.5 μm (Wong et al., 2014). These studies report 1σ variance of 9×10^{-5}, 8×10^{-5}, 6×10^{-5}, 4×10^{-4}, and 8×10^{-5}, respectively, which represent a combination of the astrophysical dayside variability of the planet, Poisson noise, and the level at which researchers could model the detector sensitivity.

2.4.2. Reanalysis of Single Eclipses

In a few cases, the *same* data have been reanalyzed and republished by different authors, and these measurements have usually differed by $< 1\sigma$: HD 189733b at 16 μm (Deming et al., 2006; Charbonneau et al., 2008), HD 149026b at 8.0 μm (Knutson et al., 2009b; Stevenson et al., 2012a), GJ 436b at 8 μm (Deming et al., 2007; Demory et al., 2007; Stevenson et al., 2010), and CoRoT-2b at 4.5 and 8.0 μm (Gillon et al., 2010; Deming et al., 2011).

Consider, however, the reanalysis of the original Harrington et al. (2007) 8 μm eclipse of HD 149026b by Knutson et al. (2009b). The latter authors found they could reproduce the original deep eclipse measurement, as well as the new, shallow depth obtained as part of thermal phase variations: "The diversity of eclipse depths (0.05%–0.09%) obtained in these fits suggests that the final result is sensitive to our specific choice of functions, fitting routines, and bad pixel trimming methods."

Finally, there are the secondary eclipses of GJ 436b (Stevenson et al., 2010) that were re-analyzed by Beaulieu et al. (2011). The latter authors found compatible values at 5.8 μm and identical values at 8.0 μm. At 3.6 μm they found that their eclipse depth depends on the reduction scheme and details of fitting, while at 4.5 μm they also favored a non-detection, but with an uncertainty 3× greater than the original authors.

2.4.3. Serial Analysis of Multiple Eclipses

For a handful of the best and brightest targets, multiple *Spitzer* eclipse observations have been obtained with the same instrument and published in *separate* papers. This is an important test of repeatability because it is semi-blind: the authors of the first paper did not benefit from knowing the result of subsequent observations (the latter authors, of course, had access both to the original and their new observations). This is in contrast to the studies listed in §2.4.1, for which researchers considered the ensemble of eclipse measurements as they fine-tuned their reduction and analysis pipeline.

The results of ten semi-blind repeatability tests are listed in Table 2.2. For each planet+waveband combination, we list the first published eclipse measurement based on a simple eclipse measurement, then a subsequent measurement obtained as part of thermal phase measurements or a multi-eclipse campaign. Note that for the HD 189733b 8 μm eclipse, the simple eclipse measurement (Charbonneau et al., 2008) was published *after* the phase+eclipse measurement of Knutson et al. (2007b), but clearly the order in which we list the measurements in no way impacts the analysis below.

For each eclipse measurement, we list the published value and uncertainty, σ. For each pair of measurements, we list the discrepancy, Δ, between the new measurement and the original. We also estimate the total published uncertainty as the quadrature sum of the first and second eclipse uncertainties: $\sigma_{\text{tot}} = \sqrt{\sigma_1^2 + \sigma_2^2}$.

Table 2.2. *Spitzer* eclipse repeat observations.

Planet	λ (μm)	Value	σ	Δ	$\sigma_{\rm tot}$	$\sqrt{\Delta^2 - \sigma_{\rm tot}^2}$	$\|\Delta\|/\sigma_{\rm tot}$	Reference
GJ 436b	8.0	5.4×10^{-4}	8.0×10^{-5}					Deming et al. (2007)
		4.52×10^{-4}	2.7×10^{-5}	-8.8×10^{-5}	8.4×10^{-5}	2.6×10^{-5}	1.0	Knutson et al. (2011)
HD 149026b	8.0	8.4×10^{-4}	1.0×10^{-4}					Harrington et al. (2007)
		4.11×10^{-4}	7.6×10^{-5}	-4.3×10^{-4}	1.3×10^{-4}	4.1×10^{-4}	3.4	Knutson et al. (2009b)
HD 189733b	3.6	2.56×10^{-3}	1.4×10^{-4}					Charbonneau et al. (2008)
		1.466×10^{-3}	4.0×10^{-5}	-1.1×10^{-3}	1.5×10^{-4}	1.1×10^{-3}	7.5	Knutson et al. (2012)
	4.5	2.14×10^{-3}	2.0×10^{-4}					Charbonneau et al. (2008)
		1.787×10^{-3}	3.8×10^{-5}	-3.5×10^{-4}	2.0×10^{-4}	2.9×10^{-4}	1.7	Knutson et al. (2012)
	8.0	3.91×10^{-3}	2.2×10^{-4}					Charbonneau et al. (2008)
		3.381×10^{-3}	5.5×10^{-5}	-5.3×10^{-4}	2.3×10^{-4}	4.8×10^{-4}	2.3	Knutson et al. (2007b)
HD 209458b	4.5	2.13×10^{-3}	1.5×10^{-4}					Knutson et al. (2008)
		1.391×10^{-3}	7.1×10^{-5}	-7.4×10^{-4}	1.7×10^{-4}	7.2×10^{-4}	4.3	Zellem et al. (2014)
	24	2.60×10^{-3}	4.6×10^{-4}					Deming et al. (2005)
		3.38×10^{-3}	2.6×10^{-4}	$+7.8 \times 10^{-4}$	5.3×10^{-4}	5.7×10^{-4}	1.5	Crossfield et al. (2012b)
WASP-12b	3.6	3.79×10^{-3}	1.3×10^{-4}					Campo et al. (2011)
		3.3×10^{-3}	4.0×10^{-4}	-4.9×10^{-4}	4.2×10^{-4}	2.5×10^{-4}	1.2	Cowan et al. (2012b)
	4.5	3.82×10^{-3}	1.9×10^{-4}					Campo et al. (2011)
		3.9×10^{-3}	3.0×10^{-4}	$+8.0 \times 10^{-5}$	3.6×10^{-4}	0	0.2	Cowan et al. (2012b)
XO-3b	4.5	1.43×10^{-3}	6.0×10^{-5}					Machalek et al. (2010)
		1.580×10^{-3}	3.6×10^{-5}	$+1.5 \times 10^{-4}$	7×10^{-5}	1.3×10^{-4}	2.1	Wong et al. (2014)

Comparing the Δ and $\sigma_{\rm tot}$ columns of Table 2.2 suggests that published eclipse uncertainties are too small: the original researchers, subsequent researchers, or both groups under-estimated the uncertainty in their measurement. Since the latter eclipse measurements are based on more data, we assume that they represent an accurate measurement and uncertainty, while the original measurements, based on a simple occultation, had under-estimated error bars.

Alternatively, the planets may be exhibiting weather that changes the eclipse depths from one epoch to the next, as predicted by Rauscher et al. (2007). Eclipse depth variability at the level of 5×10^{-4} would invalidate spectral retrieval because multi-band broadband emission spectra are constructed over a span of many planetary orbits. The weather hypothesis is ruled out in a few cases by the repeat observations discussed in §2.4.1, however.

2.4.4. Realistic Eclipse Uncertainties

We quantify the degree to which eclipse uncertainties have been under-estimated by combining Δ and $\sigma_{\rm tot}$ to obtain an empirical estimate of systematic uncertainty, following §2.1 of Lyons (1992).

In the first case, we assume there is an additional source of noise that affects single-eclipse measurements. Physically, this might correspond to how well one can model the detector given only a few hour observation of the science target. We estimate the magnitude of this systematic uncertainty by considering the distribution of $\sqrt{\Delta^2 - \sigma_{\rm tot}}$. In the one case where the epoch-to-epoch discrepancy, Δ, was smaller than the total published uncertainty, we set this quantity to zero.

The symmetric[3] distribution $[\sqrt{\Delta^2 - \sigma_{\rm tot}}] \cup [-\sqrt{\Delta^2 - \sigma_{\rm tot}}]$ has a standard deviation of $\sigma_{\rm syst} \approx 5.2 \times 10^{-4}$. We adopt $\sigma_{\rm syst} = 5 \times 10^{-4}$ for the remainder of this paper (this is somewhat greater than, but broadly consistent with, the repeatability estimate of 2×10^{-4} based on a pair of 3.6 μm transits of HD 189733b; Morello et al., 2014).

The second approach is to consider the distribution of $|\Delta|/\sigma_{\rm tot}$, which amounts to hypothesizing that single-eclipse uncertainties have been under-estimated by a constant factor. For example, researchers may under-estimate the degree to which the unknown detector model impacts eclipse depth uncertainty (numerical experiments have shown that most extant methods underestimate occultation error bars in the presence of correlated noise; Cubillos et al., 2014). The standard deviation of the symmetric distribution $[|\Delta|/\sigma_{\rm tot}] \cup [-|\Delta|/\sigma_{\rm tot}]$ is $f_{\rm syst} \approx 3.3$. We adopt $f_{\rm syst} = 3$ in the remainder of this paper.

2.5. Broadband Spectra with Empirical Uncertainties

To summarize the previous section, *Spitzer* has proven capable of photometry better than 10^{-4} and many existing eclipse measurements are likely accurate at that level: specifically, those based on multiple eclipses or taken as part of longer phase measurements (the bolded numbers in Table 2.1). Single-epoch eclipse measurements of the best and brightest targets have *not* been repeatable at this level, however. This is unfortunate because such single-eclipse measurements represent the vast majority of the broadband emission data (the unbolded numbers in Table 2.1).

Figure 2.2 shows the distribution of blackbody BIC vs. N_λ in light of empirical eclipse depth uncertainties. Values based on multiple eclipse measurements, or obtained as part

[3]The Δ-distribution is decidedly asymmetrical: researchers analyzing single-eclipse measurements have over-estimated the eclipse depth more often than not. Identifying the cause of this bias is beyond the scope of the current manuscript, so we limit ourself to properly estimating the empirical eclipse uncertainty.

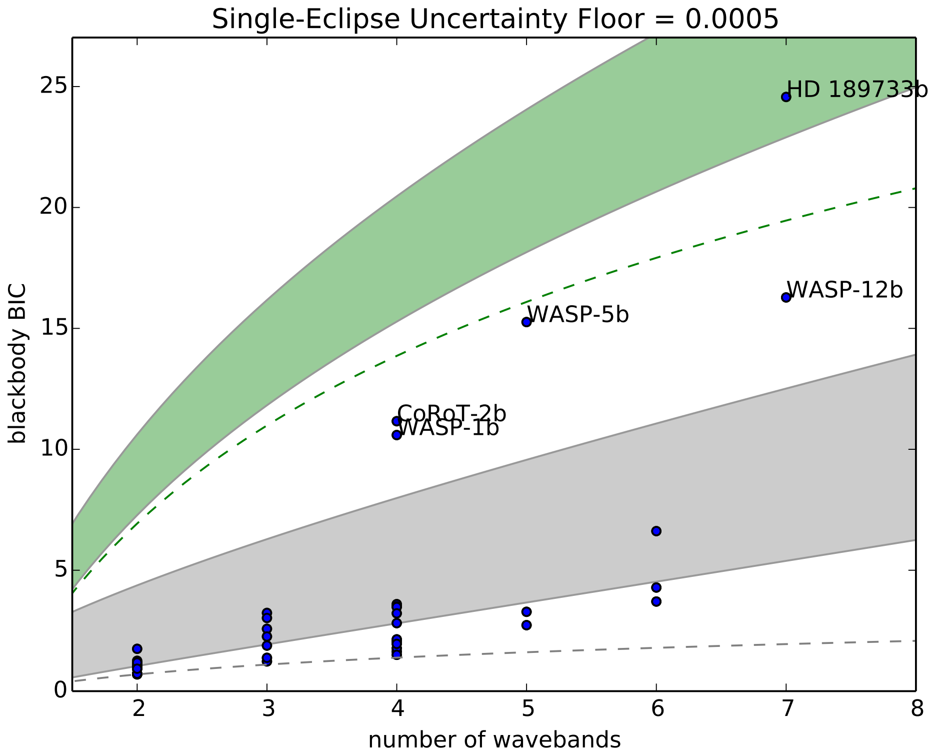


Figure 2.2. As in Figure 2.1, but we add an empirical systematic error of $\sigma_{\rm syst} = 5 \times 10^{-4}$ in quadrature to each simple-eclipse measurement. In this hypothesis, there is a floor to how precise an eclipse measurement can be without acquiring more data, so modern eclipse measurements are no more accurate than earlier attempts. Eclipse uncertainties based on multiple eclipse measurements, or an eclipse embedded in a phase measurement, are kept unchanged.

of phase measurements, are taken at face value. We add a systematic uncertainty of $\sigma_{\rm syst} = 5 \times 10^{-4}$ in quadrature to the quoted uncertainties for all single-eclipse measurements.[4] We then re-fit a blackbody and recompute the BIC for each planet using these more realistic error bars.

[4]If we had instead assumed that both the original and subsequent measurements were equally error-prone, then $\sigma_{\rm syst}$ and $f_{\rm syst}$ would be somewhat smaller, but they would have to be applied across the board, leaving our conclusions essentially unchanged.

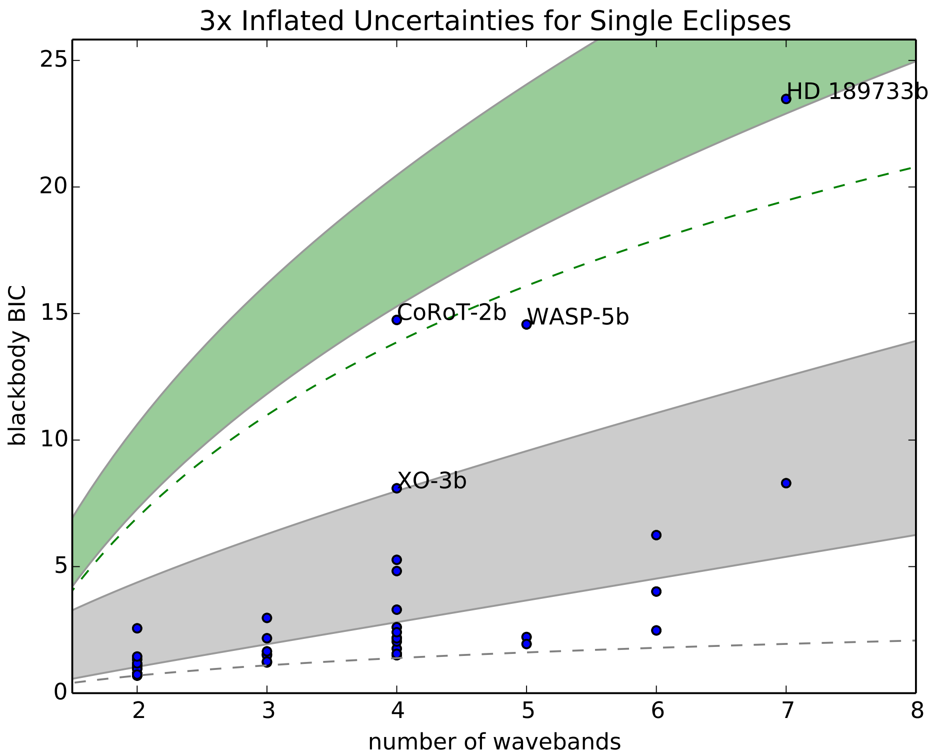


Figure 2.3. As in Figure 2.1, but we inflate the published single-eclipse uncertainties by the empirical factor $f_{\rm syst} = 3$. This scenario accounts for the possibility that modern eclipse measurements, which have much smaller quoted uncertainties than the first generation of eclipses, might really be more accurate than their predecessors. Eclipse uncertainties based on multiple eclipse measurements, or an eclipse embedded in a phase measurement, are kept unchanged.

In Figure 2.3 we inflate the published uncertainties of single-eclipse measurements by our empirically determined factor of $f_{\rm syst} = 3$. We then re-fit a blackbody and recompute the BIC for each planet using these more realistic error bars.

Under the assumption of realistic eclipse uncertainties, HD 189733b has the most featured emission spectrum and lies in the green region in both Figures 2.2 and 2.3. If spectral retrieval could achieve a perfect fit, $\chi^2_{\rm SR}/N = 0$, then it would be modestly favored as compared to the blackbody, according to the BIC. Obtaining such a good fit is not trivial for this planet because even our realistic noise hypothesis takes the published uncertainties at 3.6, 4.5, and 8.0 μm at face value.

All other planets lie at/below the dashed green line, suggesting that blackbodies are favored, *even if spectral retrieval provides a perfect fit to the data*. In any case, a researcher who has gone to the trouble of running a Markov Chain Monte Carlo to perform spectral retrieval should also estimate the evidence for their model using the posterior distribution; BIC is merely a way of approximating this. Ideally, the evidence for spectral retrieval models with different numbers of parameters could be compared using, for example, a Reversible Jump Markov Chain Monte Carlo (Green, 1995) or Nested Sampling (Skilling, 2004).

2.6. Discussion

2.6.1. The Exceptions Prove the Rule

Given the small number statistics, we expect a broad range of χ^2 values with a significant tail; the gray zone indicates the 1σ (68.3%) interval. Nonetheless, a few short period planets lie well above the gray region in Figures 2.2 and 2.3, suggesting they are poorly fit by a blackbody and hence exhibit spectral features. These features are either the hints of molecular bands, or remaining astrophysical/detector noise. The only planets that make the cut under both the $\sigma_{\rm syst}$ and $f_{\rm syst}$ hypotheses are CoRoT-2b, HD 189733b, and WASP-5b. In order to put the poorly-fitting blackbodies in perspective, we compare them to self-consistent radiative transfer models.

Self-consistent atmospheric radiative transfer models typically have between one and three tunable parameters: recirculation efficiency, optical opacity, and relative abundance of CO (e.g., Kipping & Spiegel, 2011; Deming et al., 2011) and are usually tuned by eye in order to obtain a decent fit. In what follows we will quote $\chi^2_{\rm RT}$ values from the literature (i.e., using published eclipse uncertainties). As such, the values should be compared to the blackbody BIC values shown in Figure 2.1.

As noted by Deming et al. (2011), CoRoT-2b is so poorly fit by spectral models at 4.5 μm that a blackbody fit has a smaller χ^2. In fact, Deming et al. (2011) explain the anomalous eclipse depth by invoking emission from a circumstellar accretion disk contaminating the system flux in the mid-infrared at the level of 5×10^{-3}.

Chen et al. (2014) performed spectral retrieval on WASP-5b, but the authors were unable to obtain a good fit that conserved energy, even when they allowed the atmospheric C/O ratio to vary. It is hard to imagine that a self-consistent radiative transfer model with only two variables would do any better.

The 3.6 μm photometry of HD 189733b is 5×10^{-4} discrepant from the best match 1D radiative transfer model obtained by varying two model parameters (Knutson et al., 2012). The mismatch between the predicted and measured flux at 3.6 μm contributes $(5 \times 10^{-4}/4 \times 10^{-5})^2 = 156$ to the $\chi^2_{\rm RT}$ budget, making this model a far worse fit than a simple blackbody ($\chi^2_{\rm BB} = 33$, as shown in Figure 2.1).

It is likely that bona fide fits using self-consistent radiative transfer models could provide somewhat better $\chi^2_{\rm RT}$, but this is computationally intensive and has only been performed once, to our knowledge (Kipping & Spiegel, 2011). A recent wholesale look at all extant eclipse spectra concluded that the only potentially robust area of agreement between self-consistent models and the data was the "systematic increase in the ratios to shorter wavelengths" (Burrows, 2014).

In other words, the planets poorly fit by blackbodies are also poorly fit by self-consistent radiative transfer models. The radiative transfer models could simply be wrong. There have been efforts to compare and validate exoplanet radiative transfer codes (Guillot, 2010; Shabram et al., 2011) and many have been tested against high quality observations of brown dwarfs, but it is possible that they are missing important physics relevant to irradiated planets. "Missing physics" includes atmospheric dynamics and clouds, but these are also omitted from most spectral retrieval models. We therefore hypothesize that the spectral features in extant broadband spectra are due to a combination of astrophysical + detector noise[5]; spectral retrieval provides better fits because it is under-constrained.

2.6.2. Are New Measurements More Accurate?

Most recent measurements have not yet been repeated, but one could argue that the various advances in reduction and analysis have made modern eclipse measurements more accurate than their predecessors. In hindsight, it is easy to point out poor judgements made by earlier researchers. In all cases, however, the authors were making defensible choices about how to treat the data and how to fit it. In no case has the original paper been retracted or has an erratum been published. With one exception (Beaulieu et al., 2011), researchers have only questioned the original measurements once better observations were available.

[5]The possibility that features in broadband hot Jupiter emission spectra are merely a combination of detector and astrophysical error has previously been noted by G.P. Laughlin: http://oklo.org/2013/08/21/central-limit-theorem

Researchers still make choices about their reduction scheme, and the intra-pixel sensitivity variations of *Warm Spitzer* are still modeled using the same few hours of data that are used to measure the eclipse depth. We should aspire to parametrize these choices and marginalize over them to produce accurate, if less precise, measurements. A promising avenue is to use Gaussian Processes to model the intrapixel sensitivity variations, which implicitly marginalizes over the functional form of the detector model. This strategy has been used for transit spectroscopy (Gibson et al., 2012, 2013) and to model the effect of star spots on thermal phase variations (Knutson et al., 2012).

Moreover, none of the studies reporting secondary eclipse measurements account for how the meta-parameters of reduction and analysis pipelines contribute to uncertainty in eclipse depth. At best, researchers experiment with a variety of schemes and adopt the one that minimizes the scatter in the photometry (Stevenson et al., 2012a). This amounts to optimizing the meta-parameters rather than marginalizing over them. If different choices of meta-parameters, detector parametrization, or astrophysical parametrization lead to significantly different eclipse depths (see §4.2), then one should be wary of small quoted uncertainties.

The possibility of multimodal posterior distributions should also give us pause, since neither gradient descent (e.g., Levenberg-Marquardt) nor Markov Chain Monte Carlo routines are well suited to finding global solutions under these circumstances.

In short, the current generation of single-eclipse measurements are still systematics-dominated and susceptible to many of the same problems as the previous generation. In the $\sigma_{\rm syst}$ hypothesis, there is a noise floor that affects all single-eclipse measurements, so current single-eclipse measurements are little better than the first generation. In the

$f_{\rm syst}$ hypothesis, on the other hand, the uncertainties are under-estimated by a constant *factor*, so single-eclipse measurements published today (which tend to have small quoted uncertainties) are taken to be more accurate than their predecessors. In other words, the $f_{\rm syst}$ hypothesis assumes that eclipse depth estimates are becoming more accurate with time.[6] Our results are independent of which hypothesis we choose, as discussed above.

2.6.3. Astrophysical Sources of Error

Measurement-to-measurement variance in eclipse depths is only sensitive to systematics that change from epoch to epoch: detector behavior, star spots, and exoplanet weather. There are other systematics, however, that might remain constant from epoch to epoch but that still introduce an error in our estimate of the planetary flux.

WASP-12b is the poster-child for such astrophysical sources of uncertainty, starting with the possibility of contamination from a circumstellar disk (Li et al., 2010). A change in astrophysical assumptions—namely the strength of ellipsoidal variations—affects the 4.5 μm eclipse depth of WASP-12b by 1.1×10^{-3} (Cowan et al., 2012b).

Moreover, published eclipse measurements of WASP-12b have had to be revised after the discovery of a binary companion that diluted the eclipse measurements, leading to eclipse depth increases of 8×10^{-5} to 6.5×10^{-4} in the near to mid-infrared (Crossfield et al., 2012a). In short, even if the photometry for an exoplanet system were precisely known, there is significant room for error in the dayside emission of the planet, which is the quantity we need to know for spectral fitting.

[6] It may eventually be possible to repeat this study but with so many measurements in Table 2.2 that $f_{\rm syst}$ can be a function of time, rather than constant; one could hope that $f_{\rm syst}$ tends to unity, indicating that observers are getting better at estimating the accuracy of their measurements.

2.7. Conclusions

The retrieval of atmospheric structure and composition from disk-integrated broadband photometry hinges on planets not emitting like blackbodies. We have considered the 44 short-period planets with emission measurements in multiple broadbands. If published uncertainties are taken at face value, then seven of these planets have broadband spectra that favor spectral retrieval over blackbody fits, according to the Bayesian Information Criterion—CoRoT-2b benefits the most from the additional model parameters.

In order to perform under-constrained spectral retrieval, however, it is critical to know the actual uncertainty on eclipse measurements. *Spitzer* is capable of exquisite photometry ($< 10^{-4}$), but single eclipses acquired, reduced and analyzed in isolation have only been repeatable at the 1σ level of 5×10^{-4} (or single-eclipse uncertainties have been under-estimated by a factor of 3). If one adopts such empirical uncertainties for single-eclipse measurements, then blackbody fits are preferable over spectral retrieval for all planets, with the possible exception of HD 189733b.

We conclude that statements about atmospheric composition based solely on broadband emission measurements are premature. If one adopts empirical estimate of single-eclipse accuracy, then HD 209458b and GJ 436b are well fit by blackbodies, and WASP-12b is not so poorly fit as to favor spectral retrieval. This resonates with the cautionary review of Burrows (2014). Temperature inversions and odd compositions were inferred for short period planets based on broadband emission spectra (Knutson et al., 2008, 2010; Stevenson et al., 2010; Madhusudhan et al., 2011; Madhusudhan, 2012). Our results call these claims into question. Undoubtedly, many planets have stratospheric inversions and non-solar chemistry, but there is no robust evidence for this in the current photometry of short-period planets.

2.8. Interlude I

There is a typical pattern that emerges in many parts of science, and for that matter, throughout many parts of life. It can happen over different timescales, but the basic flow is surprisingly stable: a new idea or method comes along that is enticing and pushes progressive thinking. A reaction or realization comes later, where the details of this method are scrutinized and doubt arises about its legitimacy. In the end, the tide returns to some comfortable medium, with a mature understanding of the idea that balances both its good and bad qualities. Exoplanet science is extremely new and is living through this cycle many times over right now. Amazingly we detect faraway planetary transits...are most just false positives? We find subtle hints of water in atmospheres...have we analyzed the razor-thin envelopes properly? We see unexpected eclipse depths...are our telescopes clouding the truth? This last question is particularly relevant, as we turn to tackling convoluted signals and noise in present light curves. Indeed, we have found remarkable planets with alien features and are prudent to know we know nearly nothing in the grand scheme—yet. But uncertainty alone is never good reason to stop exploring, so we now step into sensitivity variations that corrupt *Spitzer* data and vetting a popular shortcut remedy.

CHAPTER 3

Knot a Bad Idea: Testing BLISS Mapping for Spitzer Space Telescope Eclipse Observations

This chapter is adapted from Schwartz, J. C., and Cowan, N. B. 2016, Accepted in PASP.

3.1. Introduction

It is hard to characterize the atmospheres of transiting exoplanets because the atmospheric signal is 10^{-3}–10^{-5} of the stellar flux (Seager & Deming, 2010). Unfortunately, most current telescopes and instruments were not designed for these precisions.

Consider the Spitzer Space Telescope (Werner et al., 2004): many planets have been observed with its InfraRed Array Camera (IRAC; Fazio et al., 2004), and these light curves are a large part of the available data (e.g. Agol et al., 2010; Nymeyer et al., 2011; Mahtani et al., 2013; Wong et al., 2015). The pixels in IRAC are not uniformly sensitive and the target centroid (i.e. stellar position) moves on timescales of minutes to days (Ingalls et al., 2016). That means IRAC can distort the light we see (e.g. Crossfield et al., 2012b).

Many detector models have been used to deal with sensitivity variations on a pixel. Early analyses of *Spitzer* light curves used polynomials (Charbonneau et al., 2005; Knutson et al., 2008). Ballard et al. (2010, 2011) used Kernel Regression to analyze IRAC and Kepler Space Telescope data; improved versions of this method have been used by Knutson et al. (2012), Lewis et al. (2013), and Wong et al. (2015, 2016). Morello et al. (2014) used Independent Component Analysis (ICA; Waldmann, 2012) to reanalyze IRAC transit light curves. More recently, Deming et al. (2015) used Pixel-Level Decorrelation (PLD) to remove red noise from IRAC data. The authors state this method is better than modeling the sensitivity with centroids for a few reasons, including that PLD is analytically sound and runs fast.

In recent years, many researchers have used BiLinearly-Interpolated Subpixel Sensitivity mapping (BLISS hereafter; Stevenson et al., 2012a). This routine works quickly in a Markov Chain Monte Carlo (MCMC) because no *explicit* parameters are used for the detector sensitivity. Instead, BLISS divides the light curve by the current astrophysical signal at each MCMC step, averages the leftover residuals at many locations on the pixel ("knots"), then interpolates to find the sensitivity at each centroid. This means BLISS optimizes the sensitivity at each knot—it runs efficiently because the weight of each knot at the centroids' locations can be calculated ahead of time.

Many studies have used BLISS to model the intra-pixel sensitivity in *Spitzer* data, as shown in Table 3.1. Lanotte et al. (2014) and Demory et al. (2016b,a) also included the full-width half-maximum of the pixel response function in their analyses. A recent study by Ingalls et al. (2016) found that BLISS, PLD, and ICA are the most accurate and reliable ways to model IRAC sensitivity for real and synthetic observations of XO-3b. These methods can usually fit eclipse depths to within 3× the photon limit of the true values.

However, BLISS does not fit for the detector sensitivity—it merely optimizes it. The BLISS maps vary during an MCMC, but they always do so jointly with the astrophysical

Table 3.1. Works that use BLISS to model the intra-pixel sensitivity in *Spitzer* IRAC data.

Reference	**Planet/System**
Stevenson et al. (2012a)	HD 149026b
Stevenson et al. (2012b)	GJ 436
Lanotte et al. (2014)	...
Blecic et al. (2013)	WASP-14b
Cubillos et al. (2013)	WASP-8b
Blecic et al. (2014)	WASP-43b
Cubillos et al. (2014)	TrES-1
Diamond-Lowe et al. (2014)	HD 209458b
Gillon et al. (2014)	GJ 1214
Stevenson et al. (2014a)	WASP-12b
Stevenson et al. (2014b)	...
Motalebi et al. (2015)	HD 219134b
Triaud et al. (2015)	WASP-80b
Yu et al. (2015)	PTFO 8-8695 b
Demory et al. (2016b)	55 Cnc e
Demory et al. (2016a)	...
Stevenson et al. (2016a)	HAT-P-26b

model. Thus, one cannot explore the full parameter space because the BLISS map and astrophysical model are not chosen independently (Section 3.2.1). With large numbers of BLISS knots, one can also end up fitting noise in the light curve. Both of these issues mean BLISS may give astrophysical uncertainties that are too small (Hansen et al., 2014).

BLISS was introduced to side-step the computational challenge of a fully Bayesian approach (Stevenson et al., 2012a). However, nobody has tested the impact of this shortcut, nor has anybody published a rigorous study of BLISS using synthetic *Spitzer* observations, for which one knows the ground truth. Ingalls et al. (2016) tested seven techniques for removing correlated noise from IRAC data using real and synthetic observations—but only for a single hot Jupiter, XO-3b. We will therefore investigate BLISS by using a simple model of *Spitzer* IRAC light curves.

Stevenson et al. (2012a) created BLISS to handle the intra-pixel sensitivity in IRAC data because fitting $\sim 10^5$ measurements with $\sim 10^3$ model parameters in an MCMC was not feasible. This is still true, so we test light curves that have a modest number of data by using $\sim$ 25–150 BLISS knots (but see Sections 3.4.3.1 and 3.4.3.4). These sets of parameters are small enough that we can *directly* fit each knot.

We organize our work as follows: in Section 3.2.1, we describe how properly marginalizing a parameter differs from optimizing it, and use examples to show that this can affect the fits on other parameters. Then, in Sections 3.2.2 and 3.2.3, we use a toy model to show that optimizing may cause problems even with simple posteriors and Gaussian uncertainties. We describe our model of the *Spitzer* IRAC detector in Section 3.3.1, including how we make mock centroids, then introduce our astrophysical model and synthetic light curves in Section 3.3.2. In Section 3.4.1, we briefly review BLISS, and in Section 3.4.2, we compare BLISS knots and maps to the true pixel sensitivity. We then fit our light curves with MCMC and three different models for the pixel sensitivity, including two versions of BLISS, in Section 3.4.3. We discuss our results in Section 3.5 and summarize our work in Section 3.6. For those interested, the details about how we choose parameters for the pixel's sensitivity and the astrophysical signal are given in Appendices A.1 and A.2, respectively.

3.2. Optimizing Nuisance Parameters

Nuisance parameters are parts of a study that are not interesting, but have to be used to get a good answer. In the context of characterizing transiting planets, the detector sensitivity is usually modeled in terms of nuisance parameters.

3.2.1. Marginalizing vs. Optimizing

When fitting a model to data, one explores a posterior probability function: this describes how likely one's model is given each choice of parameter values. Posteriors often have many dimensions, so we show a bivariate Gaussian as a simplified example in the upper left panel of Figure 3.1. This posterior describes the arbitrary parameters X and Y, where the lighter colors show pairs of parameters that are more probable. Even though this 2D Gaussian is not oriented along X or Y, it is still highly symmetric.

Suppose now that parameter Y is a nuisance variable, and one would like the posterior (i.e. the fit) for the "interesting" parameter X alone. There are three general ways to find this, though we will focus on two for the moment. Ideally one should marginalize over Y, or integrate the 2D posterior over all possible Y-values, as shown by the (normalized) black curve in the lower left panel of Figure 3.1. Instead one could try optimizing Y, or finding the highest probability along Y for each value of X, shown in the same panel as a dashed magenta curve. For the bivariate Gaussian both methods give identical 1D posteriors on X: the median of each curve is shown with a color-coded circle, while the bars are the 1σ intervals. In other words, how one deals with this nuisance parameter Y does not affect their fit for X.

Some posteriors are less well-behaved; we show two examples in the remaining panels of Figure 3.1. The 2D posterior in the upper center is a "Gaussian butterfly," which has a narrow range of defined Y-values around X = 0 that broadens as |X| increases. The probability density varies only along X and is inversely related to the width in Y—that means the marginalized posterior for X is flat (black curve in the lower center) and the

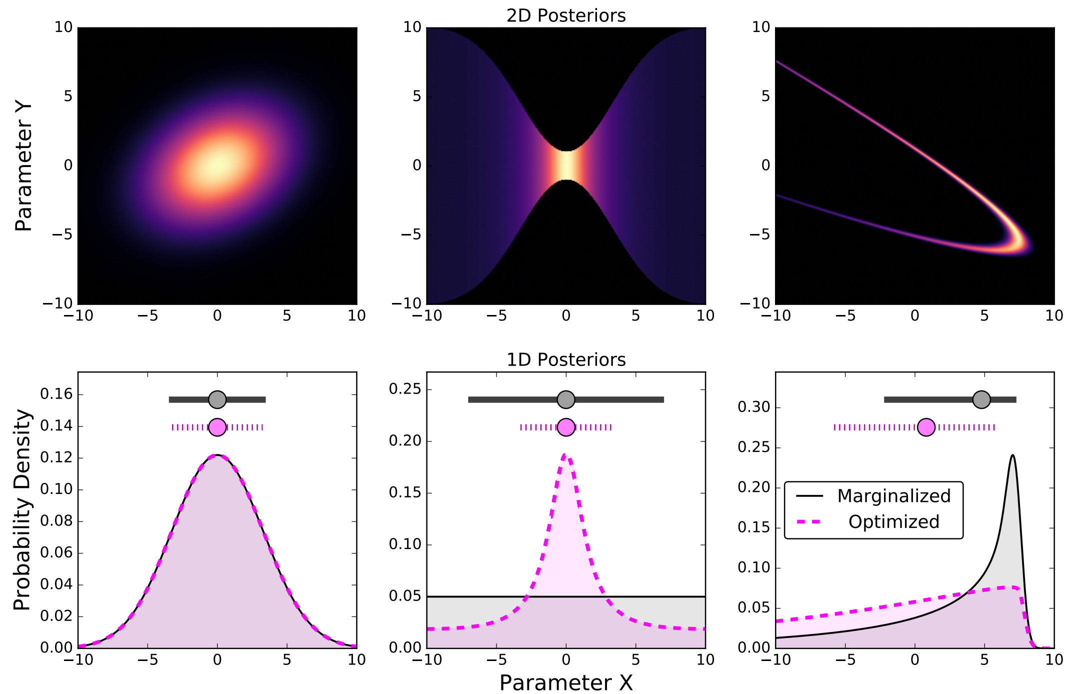


Figure 3.1. *Upper Panels:* Example 2D posteriors for the parameters X and Y: a bivariate Gaussian on the left, a “Gaussian butterfly” in the center, and a Rosenbrock banana on the right. Each color scale ranges from the maximum of the posterior (light) down to zero (dark). *Lower Panels:* The normalized 1D posteriors for each parameter X, where the black curves are the densities after marginalizing (i.e. integrating over or directly fitting) each parameter Y. Instead, one could optimize Y (i.e. find the most probable Y for each X) to get the densities shown by the dashed magenta curves. Slicing Y (i.e. cutting along the Y-value of the 2D peak) is not shown, but can be much different from optimizing Y (e.g. Rosenbrock banana), especially for high-dimensional posteriors. The color-coded circles are median values of each posterior, and the bars show 1σ intervals. For the bivariate Gaussian, one infers the same posterior and fit interval for X by marginalizing or optimizing Y—this does not happen in the other two cases. Optimizing a nuisance parameter can make the fit on another variable too precise, too conservative, or even biased.

optimized version peaks at $X = 0$ (dashed magenta curve). If one optimizes this parameter Y, their median value for X is correct (circles) but their uncertainty is too small (bars).

Alternatively, consider a 2D posterior shaped like a Rosenbrock banana function in the upper right panel of Figure 3.1. This has two thin branches that join near $(X, Y) = (7, -5)$, and the probability density does not vary the same way in both branches. The posterior for parameter X after marginalizing Y, in the lower right panel, is denser on the right and peaks around $X = 7$. By optimizing Y, though, one misses most of the banana's lower branch and so gets a flatter 1D posterior on X. In this case, the uncertainty on X is *larger* when optimizing Y, and the median is biased towards smaller X-values.

The third method we alluded to for fitting parameter X is slicing the given 2D posterior along the Y-value at its peak. This is nearly the same as optimizing Y for our first two examples, but with the Rosenbrock banana the 1D posterior for X has just two narrow, distinct peaks (not shown). For higher dimensional cases, optimizing typically falls somewhere between marginalizing and slicing the full posterior. We will return to this idea when testing BLISS in an MCMC in Section 3.4.3.3.

In general, then, optimizing parameters works well when it approximates marginalizing over those parameters: having just the silhouette of the posterior seen by the interesting variable(s) is enough to describe the nuisance parameter(s) throughout the space. This is true for the bivariate Gaussian, and in principle for multivariate Gaussians, too. Once the posterior is non-convex, has an exotic density profile, or is otherwise oddly shaped, optimizing along one or more dimensions is dicey. This may bias the best-fit values of interesting parameters and make it hard to report reasonable uncertainties.

3.2.2. Toy Model

Even if a posterior seems well-behaved, optimizing nuisance parameters can still cause problems. We demonstrate this with a toy example:

$$f(t) = (qt^2 + mt + b) + N(t; \sigma), \tag{3.1}$$

where $f(t)$ is data at time t, the q, m, and b are coefficients, and $N(t; \sigma)$ is Gaussian noise with uncertainty σ. A sample data set from this toy model is shown in the upper left panel of Figure 3.2. We use 1001 evenly-spaced times, $t \in [-10, 10]$, for a chosen set of parameters, $\{q, m, b, \sigma\}$.

The simplest way to fit these data is to use Equation 3.1, where all four parameters are fit directly. Suppose, though, that one wanted to optimize b, m, or q instead; we show examples of this strategy in the other panels of Figure 3.2. This is essentially how BLISS treats pixel sensitivity (Stevenson et al., 2012a), where detector parameters are optimized and astrophysical parameters are fitted. The idea here is to make a model with the interesting variables, then subtract this incomplete model from the data to get residuals. Then one splits the residuals into groups by time, takes the mean of each group, and finds the trend through those means. As shown, this optimizes either the offset (b), slope (m), or quadratic term (q), described ideally in Section 3.2.1. We use obvious names for each method: b-Optimize (upper right, magenta), m-Optimize (lower left, yellow), and q-Optimize (lower right, cyan).

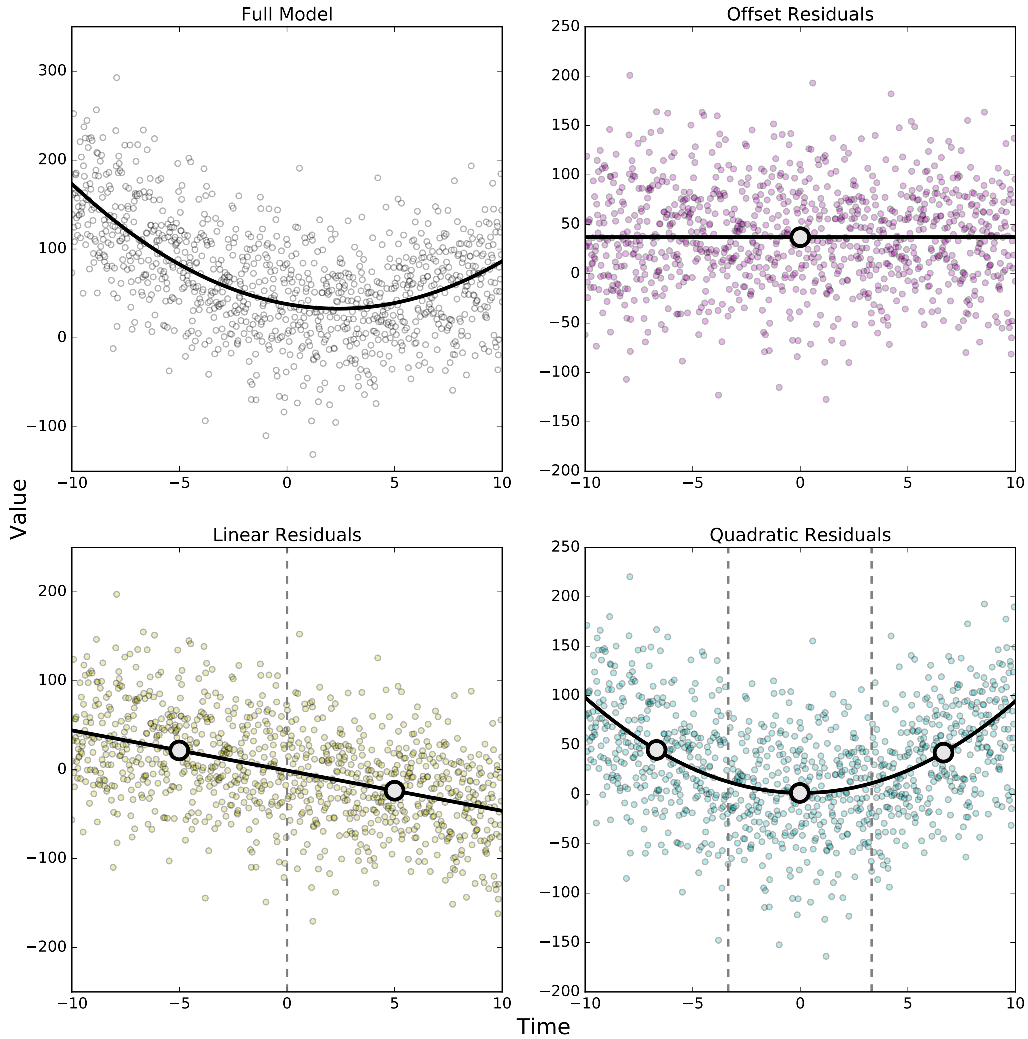


Figure 3.2. *Upper Left:* Example data generated from Equation 3.1, where the black curve is the true function without noise. *Other Panels:* Residuals left after subtracting three incomplete models, with no offset, linear, or quadratic term, from the data at upper left: b-Optimize at upper right (magenta), m-Optimize at lower left (yellow), and q-Optimize at lower right (cyan), respectively. One can estimate each missing term by splitting the residuals into time groups (dashed vertical lines), finding the mean (large gray circles) of each group, and getting the leading part of the trend (black curves) through these means. Thus, one can try to optimize each term using data residuals.

3.2.3. MCMC Fits to Toy Models

We now use the MCMC code `emcee` (Foreman-Mackey et al., 2013) to fit the data from Figure 3.2. For each of our four models, we use 240 walkers and start them in a small ball near the true parameters. We also pick uniform priors on each term in Equation 3.1. We burn-in each chain for 250 steps and run them for another 1000 steps, then thin the chains by the longest autocorrelation time, $\tau_{\max}$, that `emcee` estimates ($\tau_{\max} \approx$ 25–60 steps). Example fits are shown in the upper row of Figure 3.3. The circles are medians of each chain and bars are 1σ intervals, as in Figure 3.1.

Most of the fits to the mock data are reasonable. This is no surprise for the full model—after all, we used the same four parameters to generate the data. It is also clear that one could optimize b or m during the MCMC without hurting anything, although these schemes run no faster than the full model.

The q-Optimize method is different, though. The linear and noise terms are about the same as the other three methods, but the uncertainty on b is noticeably smaller. The center of the interval is also lower than the other methods. These walkers overlapped the same part of parameter space but tried a smaller range of offset terms.

We next try fitting 100 different data sets, where we randomly pick $q \in [-1, 1]$, $m \in [-10, 10]$, $b \in [-100, 100]$, and σ from a Normal distribution with mean 50 and width 10. We use all four methods with the same MCMC setup as before, and calculate the z-scores for each term:

$$z_\mu = \frac{\mu_\theta - \theta}{\sigma_\theta}, \tag{3.2}$$

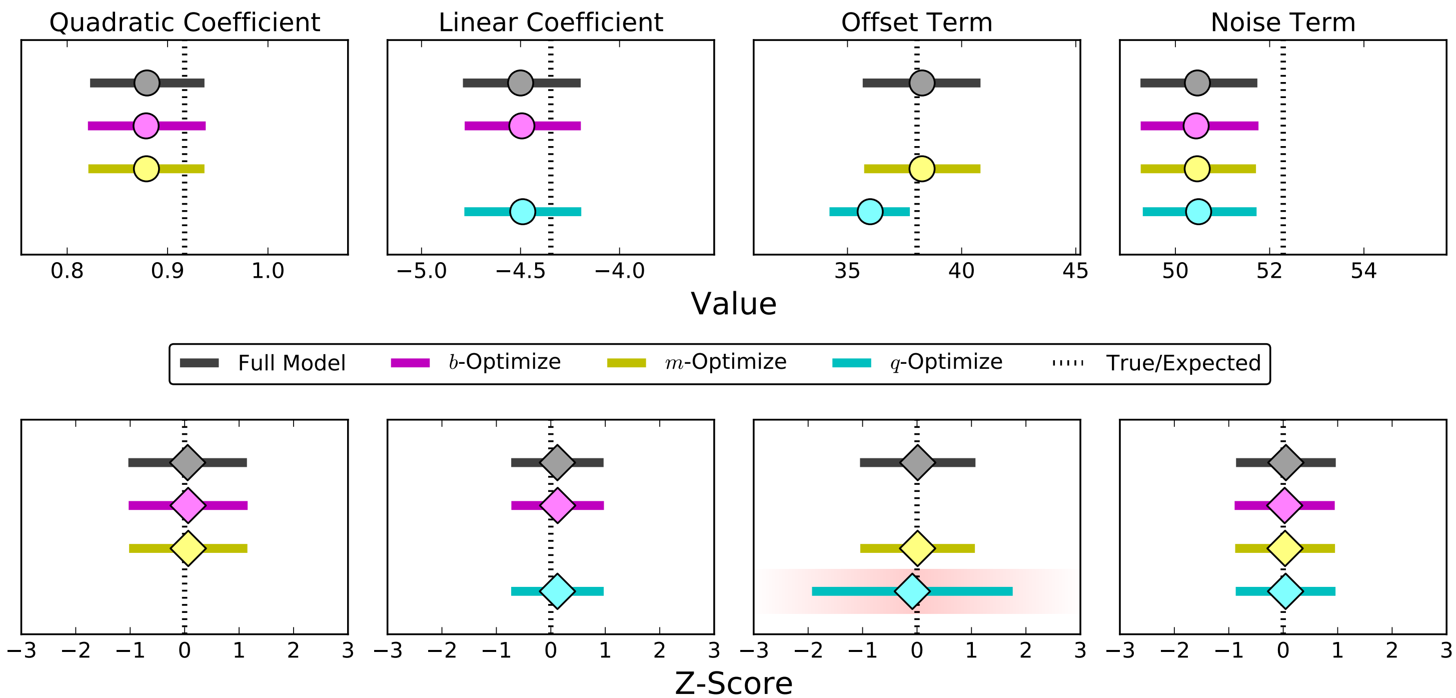


Figure 3.3. *Upper Panels:* Example fits to the data from the upper left panel of Figure 3.2, using each of the four models described in Section 3.2.2. The circles are best-fit values, bars are 1σ intervals, and the dotted black lines show the true values of each parameter. *Lower Panels:* Distribution of z-scores (Equation 3.2) for MCMC fits to 100 random data sets from Equation 3.1. The diamonds are mean values and the bars show 1σ intervals. The fit on a model term (e.g. upper panels) is unbiased if that term's z-scores are centered on zero—the uncertainty is reliable if they have a width of unity. The z-scores for the offset term in the q-Optimize model are more spread out by a factor of 2 (highlighted in red). Even with simple models and well-behaved data, optimizing a nuisance term can still lead to poor fits on interesting parameters.

where z_μ is the z-score, and $\{\mu_\theta, \sigma_\theta\}$ are the fitted value and uncertainty of parameter θ. If a parameter estimate is unbiased and accurate, then the average z-score should be close to zero and the standard deviation should be close to unity. We show the z-scores in the bottom panels of Figure 3.3, where diamonds are the mean values.

The trend in these z-scores is obvious. As we expect, the full model, b-Optimize, and m-Optimize fits look fine: on average we get close to the real parameters and have reasonable uncertainties. This is even true for parts of q-Optimize, but not the offset that this method finds. In general this uncertainty on b is too small, which is why the z-scores are more spread out than any other fit, by about a factor of 2. In other words, if one were to model this kind of data using q-Optimize, they would be too precise on their guess for b. Although this case mimics the *fits* in the lower center panel of Figure 3.1, the posterior for our toy model looks like a 4D ellipsoid (i.e. Go stone). Either q-Optimize does not "optimize" in the sense of Section 3.2.1—possible but unlikely—or the density of this posterior varies in an unexpected way.

It may seem silly to optimize the quadratic term in a quadratic equation—if one expects this term, then they should probably fit for it directly. BLISS, however, uses the same strategy to optimize the *entire* detector signal, not just one part of it. As acknowledged by Stevenson et al. (2012a), this is an expedient shortcut since fitting for $\sim 10^3$ knot values is not computationally feasible. Our example posteriors and toy model demonstrate that this shortcut may come at the price of accurate astrophysical parameters.

3.3. Synthetic Light Curves

3.3.1. Detector Model

We begin by simulating the *Spitzer* detector. Each wavelength channel of IRAC has an array of pixels, and due to the peak-up, the centroids usually stay within a single pixel for an entire eclipse observation (Ingalls et al., 2016). In real IRAC data, the image falls on different parts of the pixel because *Spitzer* both shakes and drifts slightly *and* has changes in optics due to thermal expansion and contraction.

We mimic this by modeling the centroid time-series, $\{x_0(t), y_0(t)\}$, with the pointing equations in Appendix A1 of Ingalls et al. (2016), but make two changes. We drop their short-term drift because we assume the eclipses we will model do not happen just after a re-pointing. For full-orbit phase curves where the centroids often cover larger regions of the pixel (e.g. Cowan et al., 2012b; Wong et al., 2016), including this drift could make polynomial models (Section 3.4.3) less accurate at describing the sensitivity variations. We also use regular, as opposed to fractional, Brownian motion to make the noise for their "jitter" term. This change should not influence the centroids on timescales longer than 60 seconds, i.e. the jitter period. Examples of these centroids are shown in the left panels of Figure 3.4—this observation lasts 6 hours and has 2160 data, N, or about 10 seconds per point.

The first (x_0, y_0) are both randomly chosen from $[14.7, 15.3]$ because (x, y) on the central pixel both span $[14.5, 15.5]$. We model this pixel's sensitivity using a polynomial:

$$V(x, y) = 1 + \left(\sum_{\ell=0}^{n} \sum_{m=0}^{n-\ell} c_{\ell m} (x - 15)^{\ell} (y - 15)^{m} \right)_{\ell m \neq 00}, \tag{3.3}$$

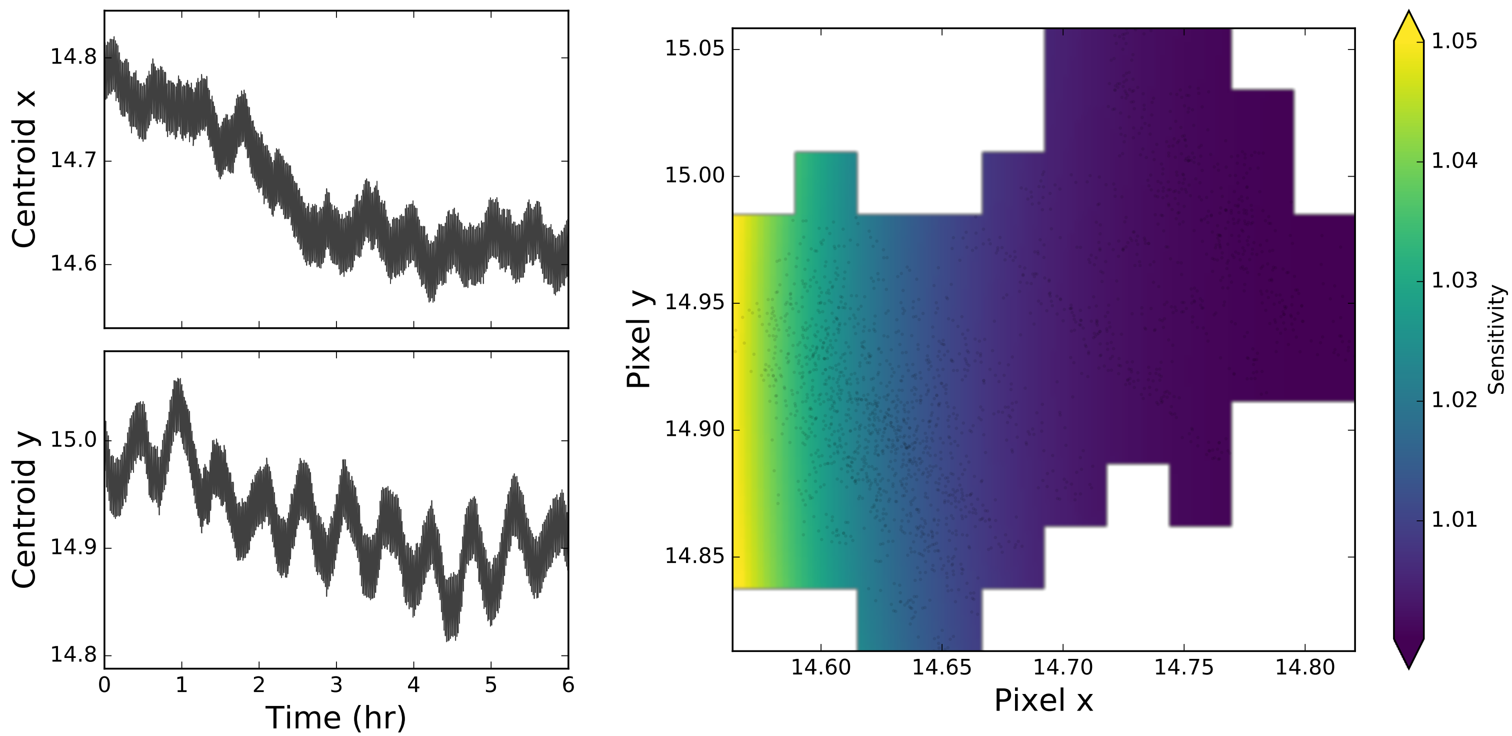


Figure 3.4. *Left Panels:* Traces of centroid position for a mock observation of a planetary eclipse. This observation is 6 hours long with $N = 2160$ measurements, or about 10 seconds per datum. *Right:* An example sensitivity map for the region of the pixel sampled by these centroids (gray dots), where lighter colors are more sensitive areas. The darkest and lightest colors are outside the sensitivity range shown—no centroids are located in these spots. We test a variety of sensitivity variations in Sections 3.4.2 and 3.4.3.4

where $V(x, y)$ is the sensitivity map and n is the polynomial order (we use $n = 7$). The $c_{\ell m}$ are coefficients, and the details about how we pick these are given in Appendix A.1. This equation keeps the average sensitivity close to unity; we show an example map in the right panel of Figure 3.4. The center of the pixel, $(x, y) \approx (15, 15)$, is the most sensitive region on the real IRAC detector (e.g. Reach et al., 2005; Cowan et al., 2012b)—this is not always true for Equation 3.3.

With the centroids and sensitivity map, we then make a detector signal, $D(t)$, using:

$$D(t) = V(x_0(t), y_0(t)), \tag{3.4}$$

that has a given amplitude, ΔD. After getting $D(t)$ and before doing anything else, we also randomly move each centroid to simulate imperfect centering. Here we use a bivariate Gaussian with standard deviations of 1% the centroid cluster's size in x and y, and a random correlation between $[-0.5, 0.5]$. These shifts are a little smaller than in Ingalls et al. (2014) and do not strongly affect our results.

Real *Spitzer* data show a variety of intra-pixel sensitivity variations in the different IRAC channels (e.g. Stevenson et al., 2012a; Triaud et al., 2015). For the example in Figure 3.4, the detector sensitivity varies about an order of magnitude more than the eclipse depth we model (below). We will test a range of other scenarios in Sections 3.4.2 and 3.4.3.4.

3.3.2. Astrophysical Model

The astrophysical signals we are interested in are planetary eclipses, and we use hot Jupiters as the model because these are the planets that BLISS is often used for. We assume our planets are on circular orbits and only consider thermal emission. Hot Jupiters exhibit thermal phase variations (e.g. Knutson et al., 2007a; Crossfield et al., 2012b; Wong et al., 2015), which we model as a sinusoid, $\Phi(t)$:

$$\Phi(t) = 1 - \alpha \cos\left(\frac{2\pi}{P_{\mathrm{orb}}} t + \phi_o\right), \tag{3.5}$$

where α is the half-amplitude, P_{orb} is the orbital period, t is the time from the start of the observation, and ϕ_o is the phase offset. The constant keeps $\Phi(t)$ close to unity, and we fix $t_{\mathrm{max}} = 6$ hrs because real observations are about that long.

Then we inject the eclipse to get the full astrophysical model, $A(t)$:

$$A(t) = \begin{cases} \overline{\{\Phi(t) - \delta_e\}}_{\mathrm{eclipse}}, & |t - t_e| \leq t_w. \\ \Phi(t), & \text{otherwise}, \end{cases} \tag{3.6}$$

where δ_e is the eclipse depth, t_e is the time at the center of eclipse, and t_w is the time from t_e to ingress or egress. We choose $t_w = 1$ hr because real eclipses of hot Jupiters usually last a couple hours. The bar in Equation 3.6 means we take the average of all data during the eclipse, so ingress and egress are instantaneous and the bottom of the eclipse is flat. The details about how we choose the other parameters for $A(t)$ are given in Appendix A.2.

Finally, we combine Equations 3.4 and 3.6 to create our model of *Spitzer* light curves:

$$F(t) = A(t)D(t) + N(t; \sigma), \tag{3.7}$$

where $F(t)$ is the flux, $D(t)$ is the detector signal from Section 3.3.1, and $N(t;\sigma)$ is photon (Gaussian) noise with uncertainty σ. We characterize our light curves using the normalized detector amplitude, $\Delta D/\delta_e \equiv \Delta D_e$, and the significance of the eclipse, $\mathbb{S}_e$, defined as:

$$\mathbb{S}_e \equiv \frac{\delta_e \sqrt{N_e}}{\sigma}, \tag{3.8}$$

where N_e is the number of data during the eclipse.

An example light curve is shown in Figure 3.5, made with the centroids and sensitivity map in Figure 3.4. The upper panel shows the astrophysical and detector signals as a dark dashed curve and an orange curve, respectively. These parts are combined in the lower panel: the brown curve is the flux one would see without photon noise, and the gray circles are data points, binned in groups of 20 for clarity. For this case, the eclipse is detected at 10σ and $D(t)$ has an amplitude $10\times$ larger than the eclipse depth. This type of detector signal is similar to IRAC data at 3.6 μm (e.g. Stevenson et al., 2012a; Cubillos et al., 2013)—we test a variety of values for ΔD_e in Sections 3.4.2 and 3.4.3.4.

3.4. Tests of BLISS

3.4.1. BLISS Method

We give a brief summary of BLISS here—for details, see Stevenson et al. (2012a). A light curve has two main parts: a detector signal (e.g. due to varying sensitivity on the pixel) and an astrophysical signal (e.g. a planetary eclipse). If one knew the astrophysical part and *divided* it out of the light curve, all that should be left in the residuals is the detector signal and photon noise.

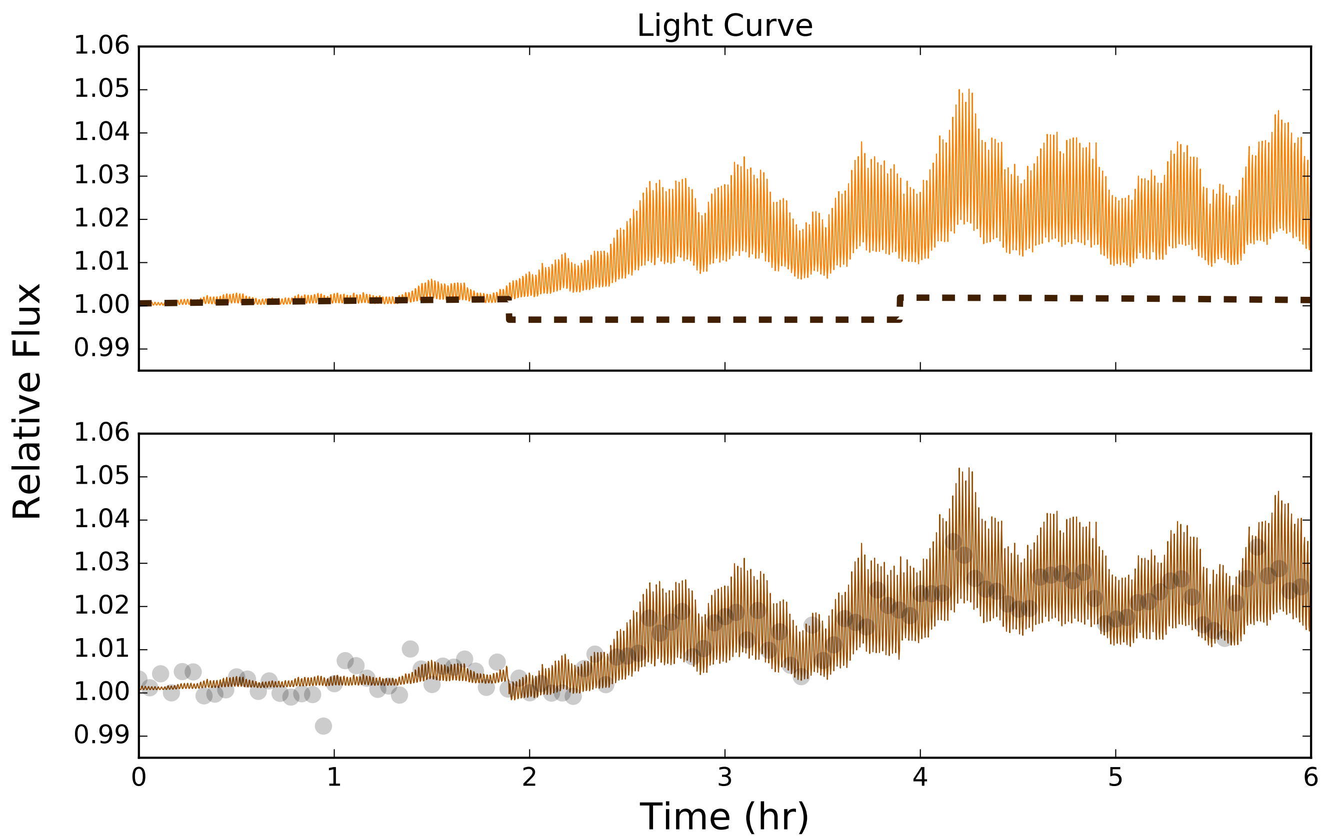


Figure 3.5. *Upper:* Examples of the two components in light curves that we model. The dark dashed curve is an astrophysical signal made with Equation 3.6. The orange curve shows a detector signal made with Equation 3.4, using the centroids and sensitivity map in Figure 3.4. The amplitude of this detector signal is 10× larger than the eclipse depth—we test other cases in Sections 3.4.2 and 3.4.3.4. *Lower:* A synthetic light curve made with Equation 3.7 and the above signals. The brown curve shows the flux with no photon noise and the gray circles are data, binned in groups of 20 for clarity. The eclipse is a 10σ detection for these data.

Each residual is paired with a centroid, so one can group the residuals with a mesh of BLISS "knots," K (left panel of Figure 3.6), take the average of each group, and set the values of the knots to these averages. This estimates what the sensitivity looks like on the pixel around the centroids, and each purple star in Figure 3.6 is a good BLISS knot, or one that has at least one centroid nearby. Other studies (e.g. Stevenson et al., 2012b; Blecic et al., 2014) often require good knots to have at least four linked centroids—those with just one nearby centroid will fit noise by definition. But, this should only affect a tiny part of the detector model and so is negligible. We explicitly try making $K = 10^2$ an ideal mesh size for our example, but this is difficult to do (Sections 3.4.3.1 and 3.4.3.4).

To figure out what $D(t)$ is, BLISS interpolates the sensitivity at each centroid by using the four surrounding knots (hence bilinear interpolation). For centroids where any of those four knots are unconstrained by the residuals (light red x-marks in Figure 3.6), BLISS does nearest neighbor interpolation (NNI) instead. Usually a few of our centroids are just outside the mesh of BLISS knots, so we extrapolate the sensitivity at those spots when we can. During the course of an MCMC, a new astrophysical signal is made at each step, the new residuals are averaged, and the detector signal is recalculated. Thus, BLISS tries to attribute unfitted variations to the detector.

3.4.2. Comparing Knots and Maps

BLISS has been used many times to handle sensitivity variations in IRAC data (e.g. Diamond-Lowe et al., 2014; Triaud et al., 2015; Stevenson et al., 2016a), and has been shown to be reliable and accurate at estimating the eclipse depths of XO-3b (Ingalls et al., 2016). But, no research has looked at the accuracy of BLISS knots or maps. We first

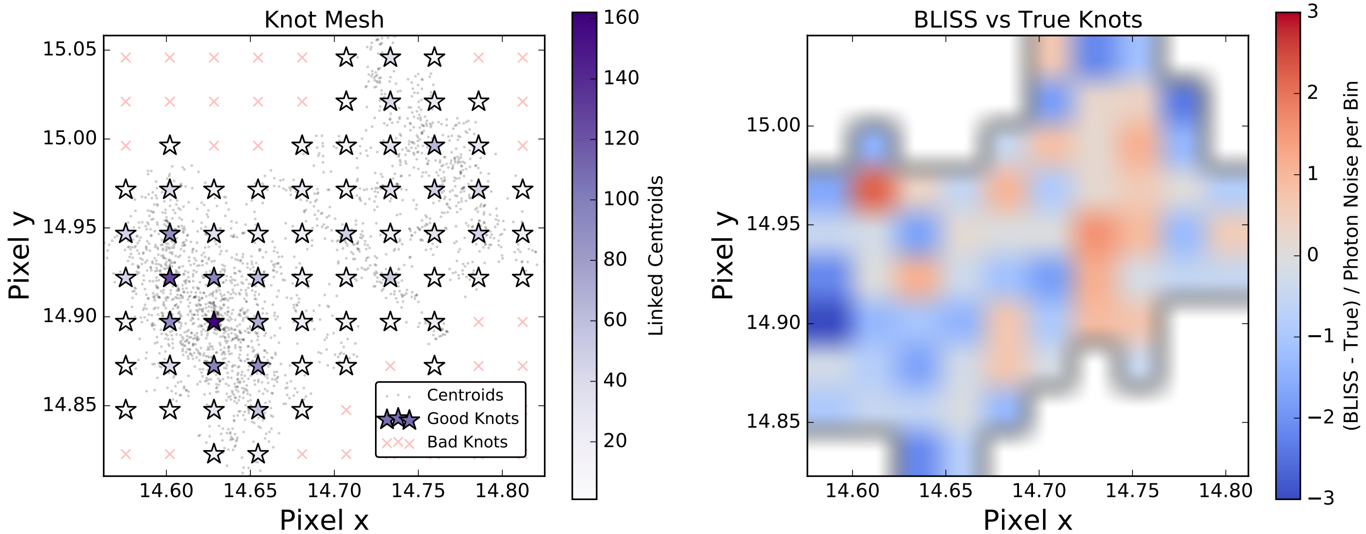


Figure 3.6. *Left:* A $K = 10^2$ mesh of BLISS knots covering the centroids (gray dots) from Figure 3.4 (these knots are chosen reasonably, but see Sections 3.4.3.1 and 3.4.3.4). The stars are good knots, or those with at least one centroid nearby, while light red x-marks are bad knots. The color scale shows how many centroids are linked with each knot; darker purple stars are knots where more data is averaged to guess the sensitivity there. BLISS then interpolates the sensitivity at each centroid using the four surrounding knots. *Right:* Discrepancies between the BLISS and true knots, using Equation 3.9. Darker reds (blues) are where a BLISS knot has a higher (lower) sensitivity than the pixel at that spot; the color scale shows up to ± 3. The average discrepancy is about -0.34, while the standard deviation is roughly 1.02—very close to the expected RMS value of unity. These knots also generate a good map (i.e. detector signal) compared to the residuals, with $\chi^2/N \approx 1.05$. BLISS knots and maps are usually accurate for the data in a light curve, but grow inaccurate when the photon noise is low.

calculate the true sensitivity at each knot's location by evaluating Equation 3.3 there—these are the values that BLISS tries to estimate.

To guess what the best-fit BLISS knots would be, we next take $F(t)/A(t)$ in Equation 3.7 and use those residuals in the BLISS routine (this estimate is good; Section 3.4.3). Then we compare the BLISS and true knot values to each other:

$$\delta k_i = \frac{(k_B - k_T)_i}{\sigma/\sqrt{N_i}}, \tag{3.9}$$

where δk_i is the discrepancy of knot i, k_B is the value of a BLISS knot, k_T is the true sensitivity at the same knot, and N_i is the number of centroids linked to that knot. The denominator in Equation 3.9 is the photon noise per bin (assuming Poisson statistics), which implicitly weights the discrepancies by the data per knot (i.e. star color in Figure 3.6). Again, δk measures how well BLISS estimates the sensitivity at the knots—we test the full map, or the interpolated detector signal, further below. We show values of δk for our example knots in the right panel of Figure 3.6. Although the astrophysical model is known perfectly here, the larger discrepancies can occur in the interior of the mesh where there are more data per knot.

The standard deviation of Equation 3.9 for all the knots tells us how reliable these estimated sensitivities are—average discrepancy matters less because *Spitzer* is poor for absolute photometry of planetary eclipses (e.g. Reach et al., 2005). Similar to z-scores in Section 3.2.3, we expect an RMS value close to unity if the knots are accurate. For example, the standard deviation on δk in Figure 3.6 is about 1.02 (average is around -0.34), so these BLISS knots are indeed a good match to this pixel's true sensitivity. We test this for other light curves by varying four parameters: the number of data points

(N), the total amount of BLISS knots (K), the eclipse significance, and the normalized detector amplitude.

We start by making an $11 \times 11 \times 50 \times 50$ logarithmically-spaced grid of $N \in [10^2, \sim 10^5]$, $K \in [5^2, 160^2]$, $\mathbb{S}_e \in [1, 100]$, and $\Delta D_e \in [0.1, 100]$, respectively. We also try a second grid where the dimensions are reversed. Then we make 3 light curves (Equation 3.7) at each grid point, get the BLISS knots as described above, and use Equation 3.9 to find the average standard deviation of δk. In general, we find a trend in RMS values with $\mathbb{S}_e \Delta D_e = \Delta D / (\sigma / \sqrt{N_e})$, which is the detector amplitude relative to the astrophysical precision on eclipse timescales (Section 3.3.2). We also find a similar trend with the average data per BLISS knot, N/K. However, the number of *good* knots for given data depends on the shape of the centroid cluster, so we focus more on $\mathbb{S}_e \Delta D_e$.

For given amounts of data and knots, when $\mathbb{S}_e \Delta D_e$ is low the photon noise is much bigger than the detector amplitude, and the standard deviation of δk is around unity. In these cases a BLISS knot is generally as accurate as the noise in the residuals. As $\mathbb{S}_e \Delta D_e$ goes up, the photon noise decreases, and the knots get closer to the true sensitivities while still being noise-limited. We eventually find an ideal regime, covering about an order of magnitude in $\mathbb{S}_e \Delta D_e$, where BLISS knots have values similar to the pixel's true sensitivity *and* the RMS of δk stays around unity. Above this range, however, the photon noise decreases so much that the standard deviation of δk grows, even though the knots stay close to their true values. These are bad levels of $\mathbb{S}_e \Delta D_e$ because BLISS is not estimating the sensitivity at the knots correctly for the expected precision. Since N/K has a similar trend, this means that when $\mathbb{S}_e \Delta D_e$ is high for given N, BLISS will estimate the sensitivity better by using more knots (i.e. smaller bins).

The full maps (i.e. sensitivity at each centroid) are comparable. When we use the BLISS and true knots to interpolate $D(t)$ for a set of centroids, both typically fit the residuals equally well (i.e. similar Chi-square values) when $\mathbb{S}_e \Delta D_e$ is in or below the ideal range. This happens in Figure 3.6, where both $D(t)$ would have $\chi^2/N \approx 1.05$. Once $\mathbb{S}_e \Delta D_e$ is $\sim 2\times$ the ideal limit for accurate knots or higher, BLISS maps usually do a little better, but both fits start to become poor. The photon noise is low in these cases, and neither map models the detector signal to within the precision of the data. On the other hand, having $N/K \sim 10$ or less means that $\chi^2/N < 1$ and the BLISS maps fit progressively more noise. Still, in most cases modeling $D(t)$ with BLISS is statistically as good as interpolating from the true sensitivity at the knots.

For example, with $N \approx 2.5 \times 10^4$ and $K \approx 30^2$, we get ideal BLISS knots when $\mathbb{S}_e \Delta D_e \in [10, 250]$ and good detector signals when $\mathbb{S}_e \Delta D_e < 500$, both roughly. We can use these values to guess how accurate the BLISS knots and maps are for the studies in Table 3.1, which often use similar N and K for eclipse observations. We estimate the detector amplitudes from uncorrected light curves or the sensitivity maps if shown, and the eclipse significances from binned light curves that have uncertainty bars. In general, we find that most studies (e.g. Cubillos et al., 2013; Stevenson et al., 2014a) have $\mathbb{S}_e \Delta D_e$ values within our ideal range—these BLISS knots and maps should be accurate. Two 3.6 μm cases to note are Blecic et al. (2013), where we estimate $\mathbb{S}_e \Delta D_e \in [250, 450]$ for WASP-14b, and Stevenson et al. (2012a), with $\mathbb{S}_e \Delta D_e \in [360, 600]$ for HD 149026b. In these studies, the sensitivity at the BLISS knots is likely starting to go bad ($1.0 \leq \mathrm{RMS}[\delta k] \leq 1.5$), but the detector signals should still be modeled well. Naturally, higher values of $\mathbb{S}_e \Delta D_e$ would be worse.

Stevenson et al. (2012a) states that, when possible, one should choose a bin size for BLISS (i.e. number of knots for given data) which does not depend on the eclipse depth and gives less scatter in the best-fit residuals than NNI. For a given light curve, it seems that one could also use $\mathbb{S}_e \Delta D_e$ (or N/K) to select an ideal number of knots for their BLISS routine. However, any of these guidelines are likely problematic (Sections 3.4.3.1 and 3.4.3.4).

We will test different sizes for the knot mesh when fitting some of our synthetic data with BLISS (Section 3.4.3.4). Though we will make practical choices for N and K to run MCMC on our light curves, these data mimic published studies and our results should apply to real *Spitzer* observations.

3.4.3. MCMC Fits to Synthetic Eclipses

We want to fit our light curves using MCMC and BLISS, but cannot use lots of BLISS knots because `emcee` would run very slowly with that many parameters. Indeed, this is why Stevenson et al. (2012a) introduced this residual optimization scheme in the first place. Instead, we start with $N = 2160$ and test for the number of BLISS knots to use, suggested by Stevenson et al. (2012a) above.

3.4.3.1. Selecting the BLISS Mesh. As stated in Section 3.3.2, the data in our main example (a 10σ eclipse with $\Delta D = 10\delta_e$; Figure 3.5) is modeled on IRAC at 3.6 μm. We therefore use Table 2 of Stevenson et al. (2012a), also for 3.6 μm data, as a guide (T2 for short). When we fit light curves like in Figure 3.5 with BLISS, the eclipse depths are usually very consistent at 1σ for $K \in [7^2, 20^2]$. Our centroid clusters are ~ 0.2 pixels wide in x and y, so these knots are spaced about $[0.03, 0.01]$ pixels apart. This matches T2 well and shows our eclipse depth does not depend on bin size.

Having the best-fit residuals be less scattered for BLISS than NNI is harder to do. One can only ensure this by explicitly fitting a light curve with both methods, not feasible for our study. Instead we approximate these fits by using $F(t)/A(t)$ from Equation 3.7 in both routines, as in Section 3.4.2. Then from T2, we compare the *ratio* of standard deviations in the best-fit residuals for BLISS and NNI, $\mathbb{R}_N^B$. We estimate that eclipse depths fit by BLISS in T2 become inconsistent when $\mathbb{R}_N^B$ drops below ≈ 0.987, at a bin size of ~ 0.06 pixels. So that our knots are spaced closer than this, we keep $K = 10^2$ as the starting mesh (i.e. BLISS bin size of ~ 0.02 pixels) and only use light curves (Section 3.4.3.4) where we estimate $\mathbb{R}_N^B \in [0.99, 1.0)$. This is our conservative attempt to have BLISS work better than NNI—true for our main light curve in Figure 3.5.

However, our choice is probably arbitrary. The value of $\mathbb{R}_N^B$ seems to depend on many aspects of a light curve, especially the detailed shape of the detector signal. Worse, when we draw new Gaussian noise in Equation 3.7 while keeping $A(t)$ and $D(t)$ fixed, $\mathbb{R}_N^B$ can be above or below unity, sometimes with equal chance. That means different photon noise *with the same uncertainty* can make BLISS look good or unnecessary for given data and K. Thus, picking the BLISS bin size, according to Stevenson et al. (2012a), can need fine-tuning.

Also, once NNI outperforms BLISS, Stevenson et al. (2012a) states that this bin size indicates the centering precision for a particular data set. But, our centroids are typically precise at about 5–15% of the bin size when $\mathbb{R}_N^B$ goes above unity. Even using the *perfect* locations of all centroids, NNI can still easily do better than BLISS—centering precision is not related to how BLISS performs. Instead, the bin size where NNI starts giving less scattered residuals than BLISS could be related to the length scale of the sensitivity variations. We hypothesize that both of the above issues can happen when fitting real observations.

Nonetheless, there are other benefits to using $K = 10^2$. Our average data per good BLISS knot is typically within $[25, 40]$. These ratios are smaller than Figure 6 of Stevenson et al. (2012a) suggests, but may be similar to other BLISS studies (e.g. Figure 5 of Blecic et al., 2013). Also, $N/K = 21.6$ and so our BLISS maps will not fit much noise (Section 3.4.2). More importantly, the estimated sensitivity at our knots should be accurate for light curves where the product of the eclipse significance and the normalized detector amplitude is less than ~ 300. The same is true for the detector signals when this product is less than ~ 600. As described in Section 3.4.2, these values of $\mathbb{S}_e \Delta D_e$ are good approximations for those in published papers. In other words, we want our fits to represent a variety of real data while remaining computationally feasible.

3.4.3.2. Models and Main Light Curve. We use three methods to handle the pixel's sensitivity. Since the true sensitivity is generated with a polynomial model, we try polynomial mapping, or P-type. Here though, we choose $n = 2$ instead of the real $n = 7$ to mimic our inexact understanding of the intrinsic detector sensitivity (we test the impact of this choice below). We also use BLISS as described by Stevenson et al. (2012a), or B-type. We further want to fit the knots directly, so we modify BLISS and make each knot a jump parameter inside the MCMC, or J-type. Everything else about BLISS is the same in the B- and J-type methods.

We use `emcee` as in Section 3.2.3, and for each method (P-, B-, and J-type) we choose the number of walkers to be $3\times$ the number of J-type parameters. The priors on all parameters are uniform, and we again start the walkers in a small ball near the true inputs. We run each chain until all parameters stabilize for at least $25\times$ the largest autocorrelation estimate, $\tau_{\rm max}$, then drop the burn-in and thin the chains by $\tau_{\rm max}$. Typically, this takes 5–20×10^3 steps and `emcee` calculates $\tau_{\rm max} \in [80, 100]$ steps. For our example light curve from Figure 3.5, we show all three posteriors on the eclipse depth in Figure 3.7. Here the real depth, $\delta_e = 5.0 \times 10^{-3}$, is shown with dashed vertical lines. At the top of each panel, we plot the median depth as a circle and the 1σ intervals with bars (as in Figures 3.1 and 3.3). Remember that here $\mathbb{S}_e = 10$ and $\Delta D_e = 10$.

All three posteriors are roughly Gaussian in shape. The J-type fit is centered near the true eclipse depth and B-type is even closer, but the latter has heavier tails and so is less precise. The P-type fit, however, peaks at $\sim 2\times$ deeper than the true value. Even though this model is the most precise, it has the worst accuracy. We find that J-type has the lowest Chi-square, 2102.9, compared to 2173.1 for P-type and 2182.1 for B-type. Note that all three models have $\chi^2/N \approx 1$. As P-type shows here, having noisy data can shift best-fit parameters away from their true values, despite Chi-square being good.

In Figure 3.8 we also compare our models to the true sensitivity projected along both axes of the pixel (as in Figure 2 of Stevenson et al., 2012a). Because we use $K = 10^2$ for BLISS to fit our main light curve, each of these projections is done with 10 bins on an axis. The true sensitivity is shown as a solid black curve, and our models have the same colors as in Figure 3.7. As expected from the Chi-square values, J-type (dotted red) matches the true variations best, though P- (dashed green) and B-type (dash-dotted blue) still follow

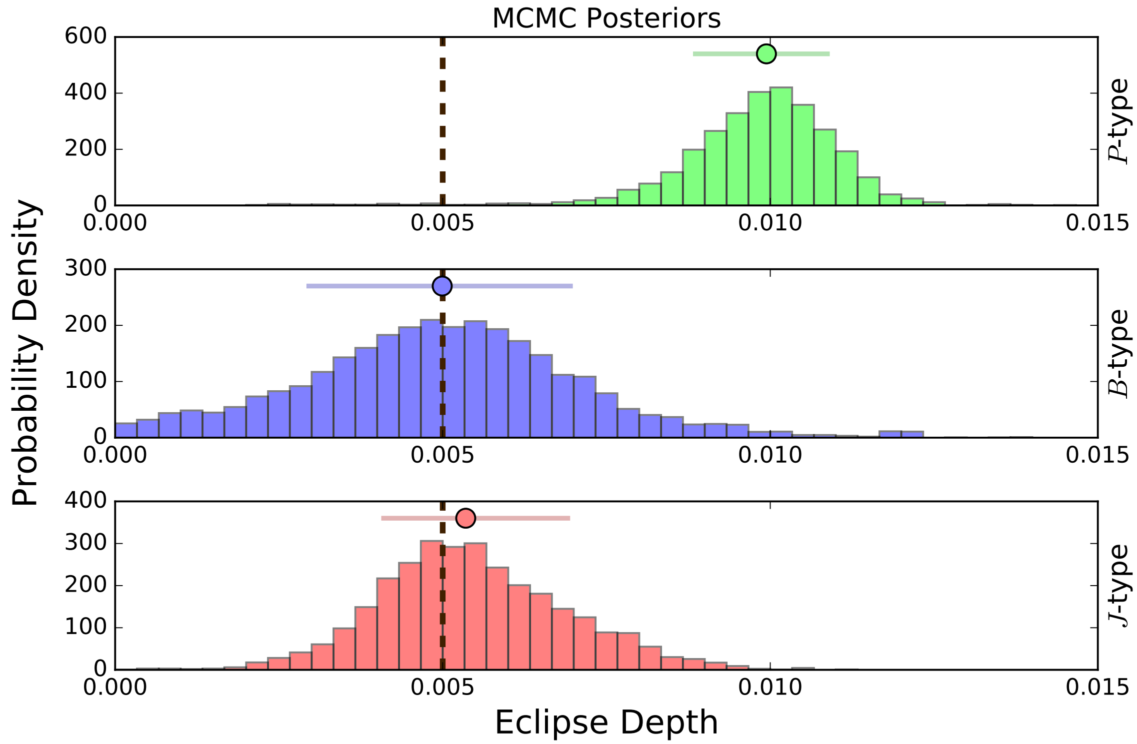


Figure 3.7. Posterior densities for the eclipse depth in Figure 3.5, fit using MCMC and three models of the sensitivity variations: polynomial or P-type at the top (green), BLISS or B-type in the middle (blue), and Jump-BLISS or J-type at the bottom (red). The dark dashed lines in each panel show the true eclipse depth, $\delta_e = 5.0 \times 10^{-3}$. The circles are medians of each posterior and the bars show the 1σ intervals. The scales for the eclipse depth are the same, though $\sim 0.5\%$ of the B-type posterior at larger depths is not shown. Our χ^2 values are 2102.9 for J-type, 2173.1 for P-type, and 2182.1 for B-type, so each model has $\chi^2/N \approx 1$. The polynomial model is the most precise method but both versions of BLISS are more accurate.

the overall patterns. The astrophysical model can balance out sensitivities here that do not match the true pixel, but this is more helpful for projections that are uniformly high or low along x or y. When we try fitting different types of light curves (Section 3.4.3.4), we find that BLISS is usually better than polynomials at matching more featured kinds of projected sensitivities.

As Stevenson et al. (2012a) describes in their Appendix A, comparing BLISS to polynomial models using the Bayesian Information Criterion (BIC; Schwarz et al., 1978) is not sound: many parameters do not overlap and each BLISS knot only interacts with a subset of the data. One could try modifying BIC to account for the latter point, but that is beyond the scope of this paper. Instead we will compare models in Section 3.4.3.4 by using accuracy and precision of the fitted eclipse depths.

3.4.3.3. Properties of BLISS. For the B-type model, we find that the best-fit BLISS knots (not shown) are mostly similar to those we estimated in Figure 3.6. The χ^2/N for $D(t)$ is higher than our original guess (≈ 1.24 versus ≈ 1.05), but both signals also look similar. We get the same results when we test other light curves, and that means we can usually estimate the best-fit BLISS knots and map well without running an MCMC. Our findings in Section 3.4.2 are therefore robust, and this supports our attempt to choose an ideal BLISS mesh in Section 3.4.3.1.

To test how the BLISS knots vary in the MCMC, we save the knots at every step in the B-type model and compare the standard deviation of each B- and J-type knot in Figure 3.9. Even though the J-type knots are free parameters, the B-types can vary more (color scale), especially those in the interior of the mesh. This is probably because there are more data per knot here, meaning the central knots have the biggest impact on the

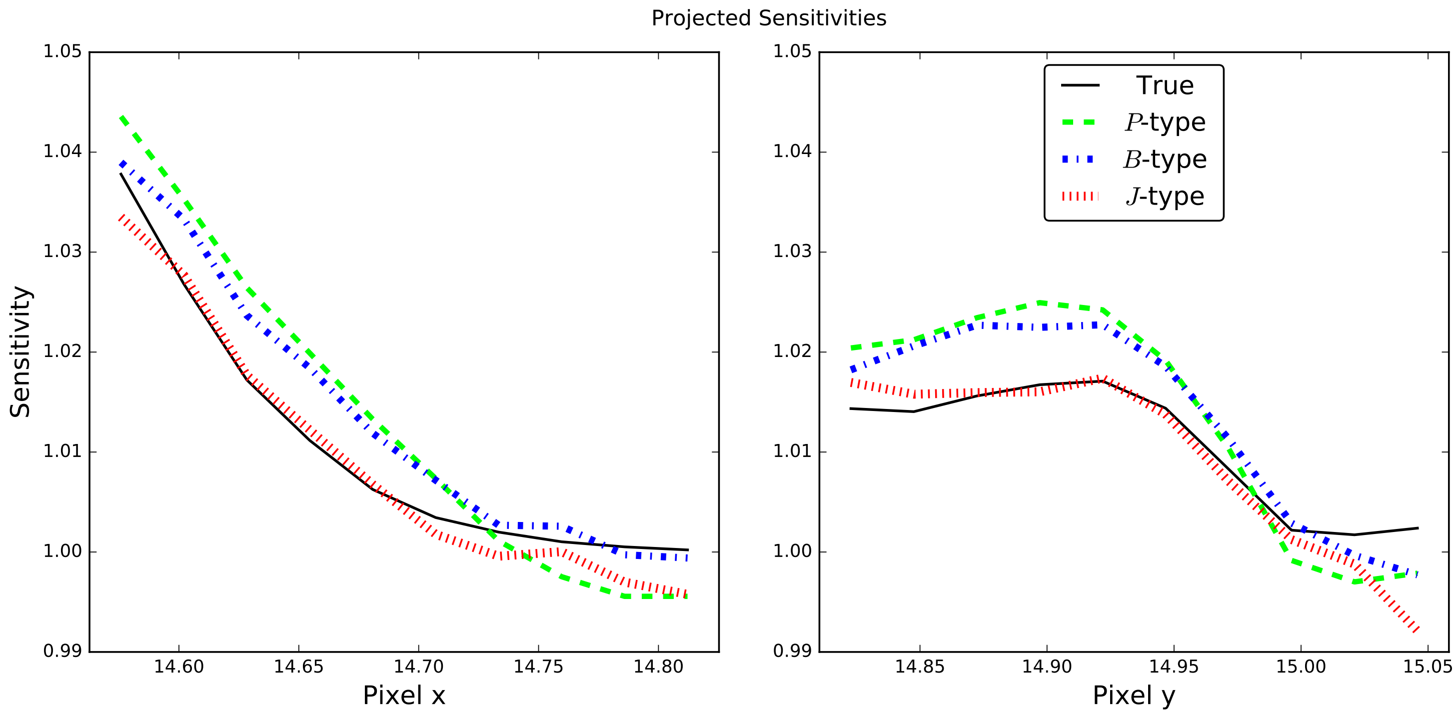


Figure 3.8. The projected sensitivities on the x- (left) and y-axis (right) of the pixel from Figure 3.4. The solid black curve is the true result, while the other curves show the best-fit models from Figure 3.7: P-type in dashed green, B-type in dash-dotted blue, and J-type in dotted red. All four projections use 10 bins along both axes. Here J-type is the model most similar to the true sensitivity variations—BLISS typically matches more featured sensitivities better than polynomials.

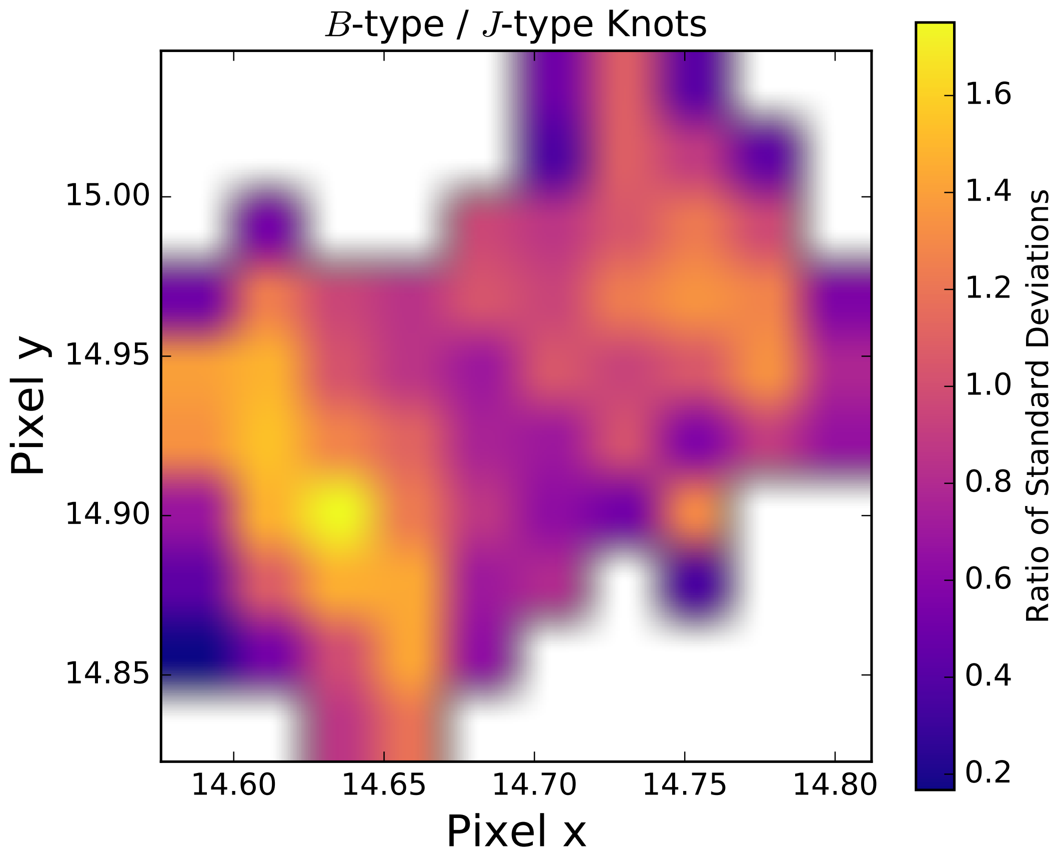


Figure 3.9. Ratio of standard deviations for knots sampled in the B- and J-type models. Lighter colors mean those knots varied relatively more in the B-type MCMC; the highest ratios are in the interior of the mesh. The value of B- and J-type knots tend to vary as much during an MCMC, regardless of whether they are jump parameters.

detector signal and so vary the most. We see this happen in every light curve we test (i.e. ratio of standard deviations between $[0.3, 2.0]$ typically), so in general BLISS knots act like real variables rather than fixed parameters.

We also try slicing through the J-type MCMC chain (i.e. posterior; Section 3.2.1) in the knot parameters. This shows how the fit to the eclipse depth changes when fixing the knot values, which should be a worst-case scenario for BLISS. With 65 good knots, though, this is tricky. For example, the density of a ν-dimensional Gaussian depends on the σ-scaled distance d from the mean (Mahalanobis distance; e.g. De Maesschalck et al., 2000), where d^2 has a χ^2_ν distribution (e.g. Tong, 2012). The chance that a point lies within $d = 1$ in our case, or 1σ in *all* knots, can be estimated as $\mathrm{CDF}(1^2; \chi^2_{65}) \sim 10^{-48}$. Our thinned chains only have $\sim 10^4$ samples, so it is near-impossible to have any sample close to the best-fit value of every knot. Discrete samples are often *very* spread out in a high-dimensional space.

In practice we take slices much larger than 1σ through the J-type knots to capture close to 10% of the samples. The above example predicts this happens when $d \approx 7.13$. When we slice around the maximum likelihood value of the knots, we only need $d \approx 2.42$. The median eclipse depth is about 4% higher than in the full chain, and the interval in nearly unchanged. If we slice around the median knot values instead, we only need $d \approx 2.04$. The median eclipse depth goes down by $\sim 5\%$, but the interval is now about 18% smaller. The low d-values we find imply that the J-type posterior is not a multivariate Gaussian.

When we test other light curves, the J-type slices often look similar. Median eclipse depths are usually within 10% of those in the full chains (i.e. good for 10σ eclipses or better), and the intervals between 20% narrower to 10% *wider*. The exceptions are when the eclipse significance is low: these fit intervals on the depth are around half the width of those in the full chains. But generally, slicing through the J-type posterior—which limits the value of each knot—does not affect the fitted eclipse depth much.

3.4.3.4. Varying the Data and BLISS. So far we have (mostly) considered the fits for a single light curve. We now try fitting different data sets and changing how many BLISS knots we use. For consistency, we fix all eclipse depths to $\delta_e = 5.0 \times 10^{-3}$ and test 5 light curves, or 10 where stated, per case we consider. We randomly generate these synthetic data, but visually inspect them to make sure the detector signal is not mostly flat, which happens about 10–20% of the time. Note that we only explicitly try to have BLISS work better than NNI (Section 3.4.3.1) in our main type of light curves (circles below; includes Figure 3.7) and when we later modify the sensitivity variations for this type. Other cases are experiments on changing some aspect of the data or BLISS.

Because we find above that the B- and J-type models are similar, we drop J-type from here on to speed up our fits. We repeat the P- and B-type MCMCs as described in Section 3.4.3.2, and since there are several changes to consider, we split these sets of light curves into groups. We find all $\chi^2/N \sim 1$ and either model can have the lowest value unless stated otherwise. Bear in mind that the following figures only show about a third of our MCMC fits—we have tried other (sometimes uninteresting) parts of the parameter space.

We first try varying both the eclipse significance and normalized detector amplitude, and show the mean and standard deviation of the z-scores (Equation 3.2) for the eclipse depth in the left panel of Figure 3.10. The P-type models are colored green and the B-types are blue. There are also two kinds of z-scores: the darker markers are the fits, while the lighter markers use more conservative intervals we get by testing for time correlations in the best-fit residuals (β plots; e.g. Pont et al., 2006; Cowan et al., 2012b). These pairs are clarified with connecting lines and the lighter markers are only shown when they do not overlap the darker version.

In Section 3.2.3 we described that parameter fits are reliable if, after many samples, the z-scores on the fits have an average of about zero and a standard deviation around unity. Each marker here only uses 5 samples, so the background shows the scatter we get when drawing, via Monte Carlo, 10^7 sets of 5 samples from a standard normal distribution. Lighter areas are more probable and the dashed magenta ellipse contains 99% of the Monte Carlo sets. If a marker is outside this region, it likely means the eclipse depths in that case are being fit unreliably. BLISS has some trouble when the eclipse is noisier and the detector signal is larger (blue diamond). The polynomial model has suspect fits when $\mathbb{S}_e \Delta D_e \geq 100$, with the darker green star outside the plot at about $(7.0, 4.7)$. Including β factors makes the z-scores reasonable for the higher two cases (lighter green square and star).

Z-scores combine the accuracy (i.e. discrepancy from a true value; numerator) and precision (i.e. width of an interval; denominator) of each *individual* fit. By separating these pieces, we can also compare the overall accuracy and precision for *types* of fits. The right panel of Figure 3.10 plots the reciprocal of both the median discrepancy and median interval width—accuracy goes up logarithmically towards the top and precision towards the right. The uncertainty bars show the interquartile ranges when these are larger than the size of the markers. Because we test 5 samples in each case, this means the uncertainty bars ignore the single highest and lowest accuracy and precision we find. Ideally markers will be on or near the solid black line, where accuracy equals precision and the fits have

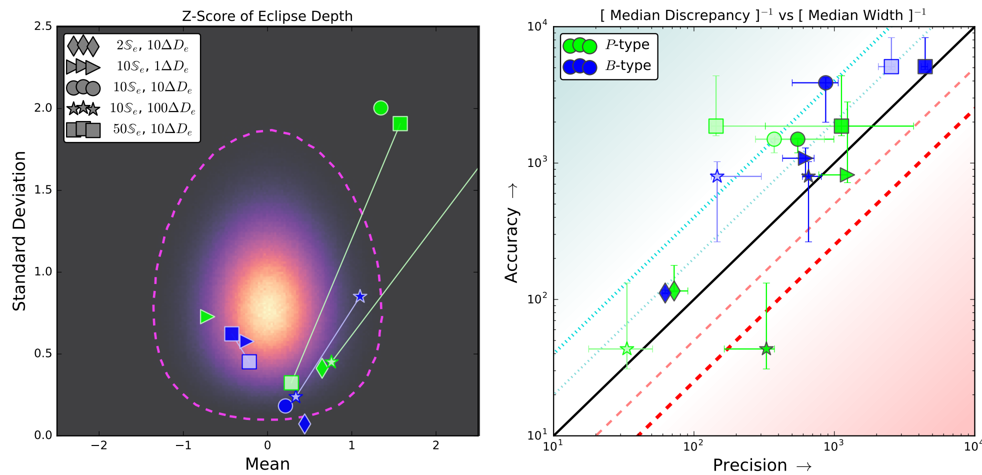


Figure 3.10. *Left:* Mean values and standard deviations of the z-scores obtained by fitting the eclipse depth in different types of light curves; here we vary the eclipse significance and normalized detector amplitude. Each marker uses 5 samples—the fits from Figure 3.7 are part of the circles—with the P-type data shown in green and the B-type in blue. Lighter markers show, when significant, how the z-scores change by accounting for time-correlated residuals via the β method (Section 3.4.3.4), where the connecting lines are for clarity. The original green star is outside the plot at roughly $(7.0, 4.7)$. The background shows the expected scatter (10^7 Monte Carlo) for sets of 5 reliable z-scores, where lighter colors are more probable and the dashed magenta line is the 99% ellipse. *Right:* The reciprocal of both the median discrepancy (i.e. accuracy; z-score numerators) and median width of the fit interval (i.e. precision; z-score denominators) for these eclipse depths. Both axes are logarithmic. The bars show the interquartile ranges in accuracy and precision. Note that inflating uncertainties on the eclipse depth via the β method (lighter markers) only affects the precision of the fits. The solid black line shows ideal cases where accuracy and precision are equal (i.e. maximum predictive power). Parallel, the dotted cyan and dashed red lines show where the ratio is $2\times$ and $4\times$ too conservative and too precise, respectively—the cyan and red background shows how these trends continue. BLISS gives better fits in our main case (circles) and when the eclipse significance is higher (squares), but the polynomial model can also do well (e.g. triangles).

maximum predictive power. If both models are on this line, the one closer to the upper right corner is preferred.

Towards the upper left the fitted eclipse depths are too conservative. The dotted cyan lines show where the accuracy is $2\times$ and $4\times$ larger than the precision (e.g. green square and blue circle). Worse, in the other direction the fits are too confident, the dashed red lines showing where accuracy is $2\times$ and $4\times$ *smaller* than precision. The green star is outside the latter line, but similar to the left panel, accounting for time-correlated residuals (i.e. inflating the uncertainties on the eclipse depth) moves this marker close to the ideal ratio. The blue square and star have reasonable z-scores but are moved off the maximum predictive line by β factors—we will return to this point when testing other sensitivity variations later on. These β factors can only decrease the precision of the fits; they cannot affect the accuracy.

Relative to our main example (circles), increasing the eclipse significance (squares) helps BLISS more than the polynomial model. In fact, B-type is preferred in both these cases due to, respectively, more reliable z-scores or better accuracy and precision. We find P-type is the preferred model for the lowest value of $\mathbb{S}_e \Delta D_e$ (diamonds). For the highest value (stars), the polynomial fits have more predictive power, but are less accurate *and* precise than BLISS. Unexpectedly, P-type is at least as precise as B-type in three of these five cases (diamonds, triangles, and circles). This is unusual because BLISS has been shown to perform better than a second-order polynomial on real *Spitzer* data (e.g. Stevenson et al., 2012a; Blecic et al., 2013). BLISS is the more precise model when $\mathbb{S}_e \Delta D_e \geq 500$, though, especially after including β factors. Since these values are at or above the limit of $\mathbb{S}_e \Delta D_e$ for accurate BLISS maps (Section 3.4.3.1), it is not surprising that β factors decrease the precision of these fits. But again, the z-scores for the B-type square and star are reasonable to start.

We show similar z-score, accuracy, and precision data as Figure 3.10 for all of the remaining figures. In Figure 3.11 we test how the fits change with the number of data or BLISS knots, while keeping $\mathbb{S}_e = 10$ and $\Delta D_e = 10$. The blue triangle, circle, star, and pentagon use the same 10 light curves, and since changing the BLISS mesh does not affect P-type, these fits should be compared to the green circle. The z-scores for both models are acceptable (i.e. markers inside the dashed magenta ellipse) in all new cases after using β factors for the green diamond. Note that the blue diamond is behind the green circle in the right panel. In every case the polynomial model and BLISS overlap in precision, given the uncertainty bars.

We see weak trends for both models when varying the amount of data: P-type increases in precision yet gets a little less accurate, and B-type increases slightly in accuracy. Actually, the polynomial model is preferred when we use more data (squares). Changing the number of BLISS knots affects B-type in the right panel, but all four cases mutually overlap in precision and even accuracy. We choose these light curves so that $K = 10^2$ (blue circle) should be optimal for BLISS (Section 3.4.3.1). However, according to Stevenson et al. (2012a), the eclipse depths we fit should not depend on the number of knots, unless we make K much smaller. Figure 3.11 confirms this.

When we also fit some light curves from these cases using NNI (not shown), the best-fit residuals for any K are always less scattered than for BLISS. This does not happen when the eclipse significance is extremely high, or in samples we test from Figure 3.10 with $\mathbb{S}_e = 50$ and $K = 10^2$. Yet here, the accuracy and consistency of our B-type fits are the same *or better* when compared to NNI. Therefore, our method in Section 3.4.3.1 to properly select K for BLISS may have issues (e.g. $\mathbb{R}_N^B$ must be lower), or the criteria in Stevenson et al. (2012a) may not work in general. Both ideas could be true.

Next we test how having more red noise can affect the fits. In Figure 3.12 we multiply an extra noise (Brownian) into the light curve to mimic different kinds of time-correlated features (i.e. other than intra-pixel sensitivity variations). We use the same 5 light curves and red noises in each case, meaning we only change the amplitude (relative to the mean) of the noises, not their structure. As often before, P- and B-type have similar precisions every time.

The case with highest noise, at $5\times$ the detector amplitude (stars), is clearly bad. The z-scores for P- and B-type are far outside the plot even with β factors included ($\sim$ 10–35 on both axes), and the fits are very over-precise. At $1\times$ the detector amplitude (squares), β factors move BLISS close to intersecting the maximum predictive line in the right panel, but not inside the 99% ellipse in the left panel. We find other cases where this outcome is more pronounced, which is a good lesson: the accuracy and precision of a model are separate scalars. Z-scores are a discrepancy *paired* with a fit interval—those specific pairings matter. That means different sets of z-scores can give the same accuracy and precision. Just because a model does well on average does not mean the individual fits are reliable, and vice versa.

The case with extra noise at $\frac{1}{5} \times \Delta D$ is curious (triangles). The z-scores for B-type are reasonable, and those for P-type are acceptable when using β factors. Moreover, both models have near-ideal accuracy and precision. Thus, adding a low amount of time-correlated noise to the synthetic data actually *improves* the predictive power of both

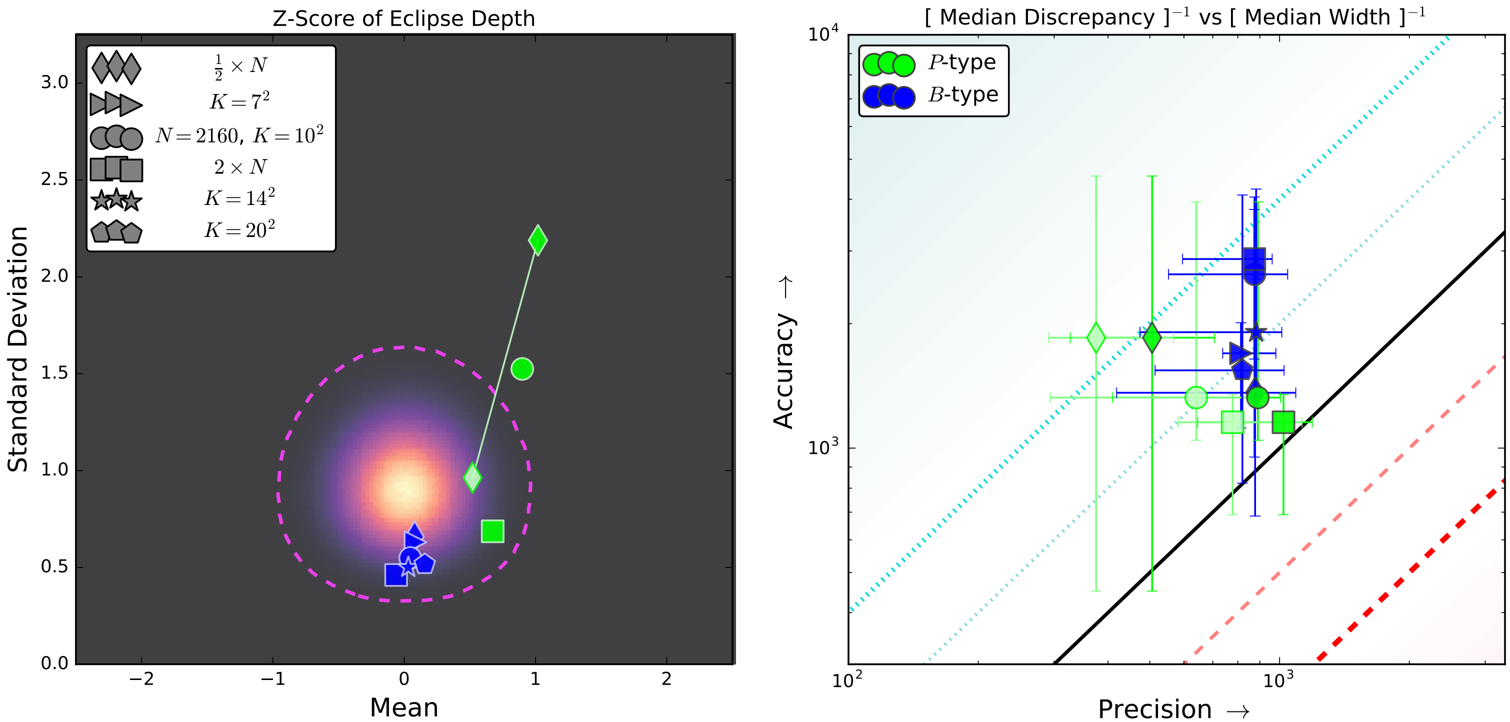


Figure 3.11. Z-scores, accuracy, and precision of fitted eclipse depths when we vary the amount of data (N) or the total number of BLISS knots (K). We test 10 samples (instead of 5) in each case here—the background and dashed magenta ellipse in the left panel account for this. The circles include the fits from Figure 3.10; note that the green circle covers the blue diamond in the right panel. Here the green circle should be compared with the blue triangle, circle, star, and pentagon, which all use the same light curves. When increasing the amount of data (diamonds to circles to squares), BLISS increases in accuracy and the polynomial model moves towards the maximum predictive line (i.e. solid black), though these trends are weak. We confirm that the accuracy and precision of BLISS are consistent when using different numbers of knots, as expected from Stevenson et al. (2012a) and Section 3.4.3.1.

fits—especially for BLISS which has insignificant β factors. This bodes well for fitting eclipse depths in real light curves because it suggests that one may not need to perfectly model every source of red noise.

We further test what happens to the fits when we modify the sensitivity variations on the part of the pixel under the centroids. Note that we already place the centroids at many locations on the pixel (Section 3.3.1) to have different terms in Equation 3.3 dominate the detector signals (Appendix A.1). In Figure 3.13 we compare two forms of the pixel's actual sensitivity variations, using two combinations of $\mathbb{S}_e$ and ΔD_e. The circles and squares (taken from Figure 3.10) have light curves made with "P-like" variations, or the polynomial $V(x,y)$ in Equation 3.3. The stars and diamonds use "B-like" variations on the pixel instead. For these we define the sensitivity as random Gaussian values at the locations of the BLISS knots. Then we interpolate the sensitivity between these spots using bivariate splines, similar to how the BLISS routine maps $D(t)$ at the centroids.

At first glance BLISS looks like the perfect model for the B-like scenario, but it is not. Remember, BLISS estimates the sensitivity at a knot by averaging the residuals (i.e. flux divided by an astrophysical model) at centroids in a bin around that knot. If that bin contains a local peak or valley in the variations, this can throw off the estimate the closer that feature is to the knot. Even when the knot values are accurate, interpolating $D(t)$ well is tricky when the pixel's sensitivity has small-scale structure, especially because the number of knots cannot be increased arbitrarily (Sections 3.4.2 and 3.4.3.1). BLISS interpolates linearly between knots adjacent in x or y, so we postulate that the routine could only *exactly* match sensitivities that vary like a plane across the pixel. Unfortunately

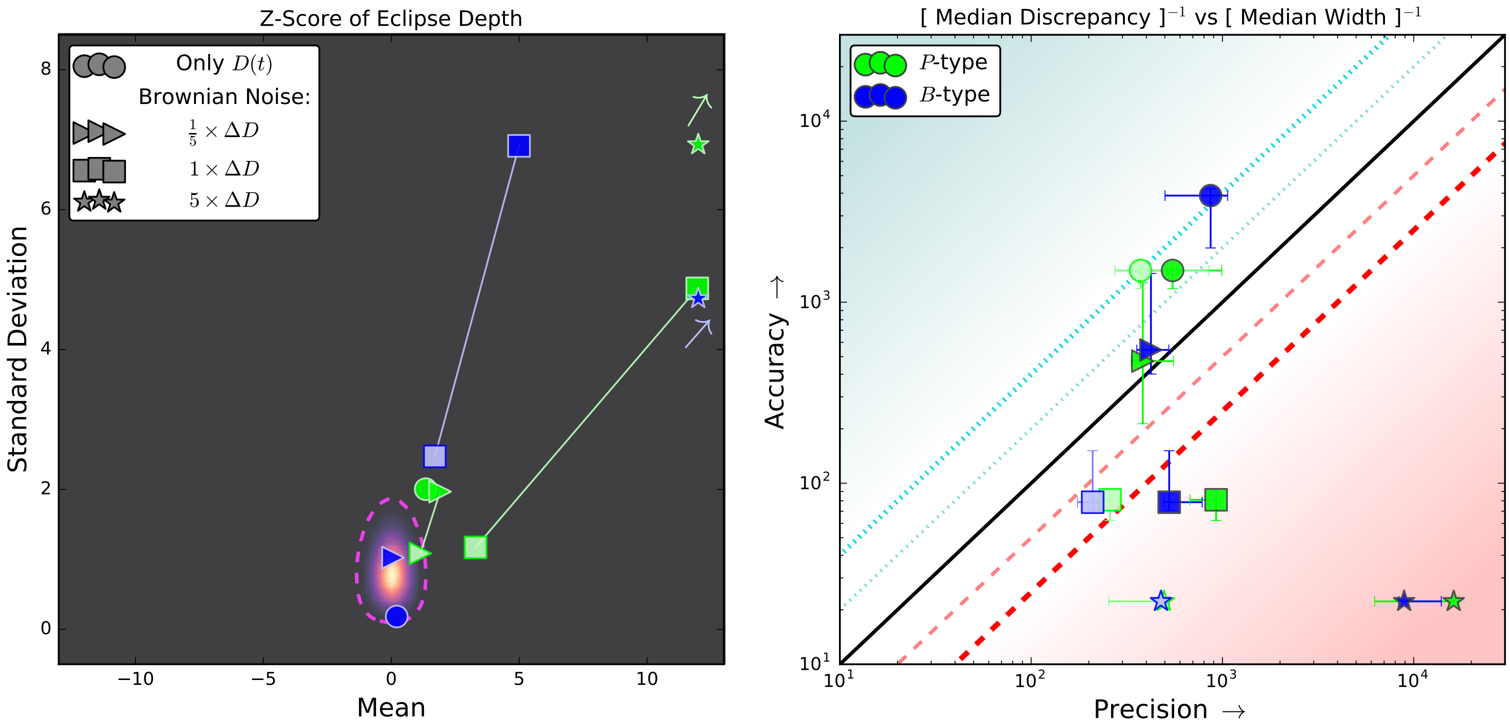


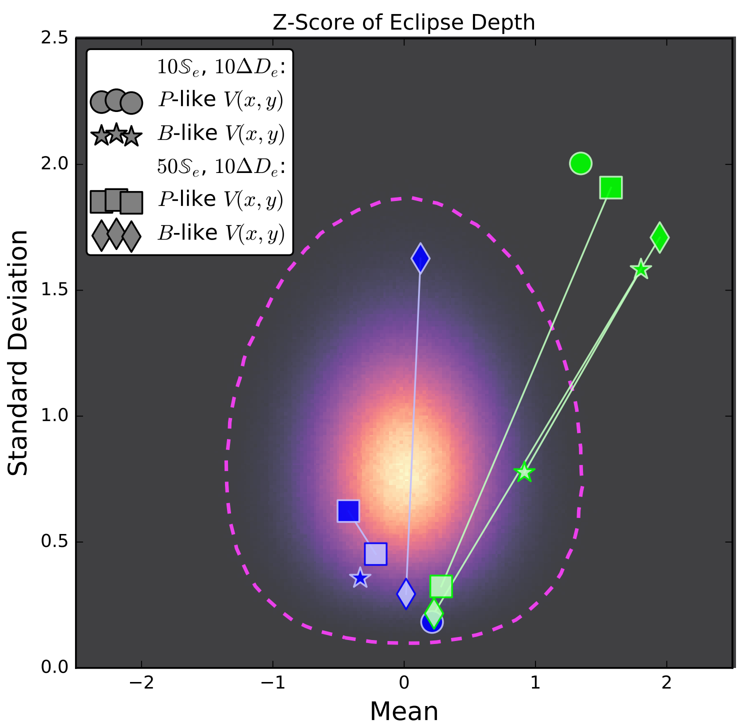
Figure 3.12. Z-scores, accuracy, and precision of fitted eclipse depths when we multiply different levels of red (i.e. Brownian) noise into the light curve, in terms of the detector amplitude. The circles are taken from Figure 3.10. All cases use the same 5 light curves and red noises—we only change the amplitudes of the latter relative to their means. Both models are poor in the case with highest noise (stars), where the arrows show that the z-scores are far outside the plot. In the moderate case (squares), the B-type fits approach the maximum predictive line with β factors included (right), even though these z-scores are outside the 99% ellipse (left). This means a model that fits eclipse depths well on average may still be unreliable. In all cases here, the polynomial model is as precise as BLISS. Interestingly, having a low amount of red noise in a light curve (triangles) may actually improve the predictive power of both models, especially BLISS.

polynomials would also fit exactly in these cases. Nonetheless, BLISS should handle our B-like variations better than polynomial models.

When $\mathbb{S}_e = 10$ and $\Delta D_e = 10$, both models decrease in accuracy when switching to the B-like scenario (circles to stars). We also find that β factors are important for the polynomial: it becomes the better predictive model despite BLISS having higher precision. Even stranger is that P-type always has a lower χ^2 than B-type. When $\mathbb{S}_e = 50$ and $\Delta D_e = 10$, having B-like variations means the median fit for BLISS moves more than for the polynomial model (squares to diamonds). However, with or without β factors, BLISS is more precise than P-type and is the preferred model.

The blue diamond here is similar to cases from Figure 3.10. Despite good z-scores, the β factors for this model change the fits from ideal to very conservative (i.e. precisions $\sim$ 3–10$\times$ less than accuracies). This is different from Figure 3.12, where the poor fits show up in the z-scores. In other words, both panels of Figures 3.10–3.13 are important to see how well a model fits a certain type of light curve. Since we find that β factors can penalize poor and reasonable fits just as much, using them to tune one's precision is not always wise.

To summarize Figures 3.10–3.13, the eclipse significance and detector amplitude affect the precision and accuracy of a fitted eclipse depth. Changing the amount of data by factors of two weakly affects the accuracy of BLISS and both the accuracy and precision of polynomial models. Using different numbers of BLISS knots gives consistent eclipse depths (as expected), but we find that heuristics for choosing the bin size—here and in Stevenson et al. (2012a)—are questionable. Large amounts of red noise in a light curve are bad, but having low levels can in fact improve both models' fits, particularly BLISS.



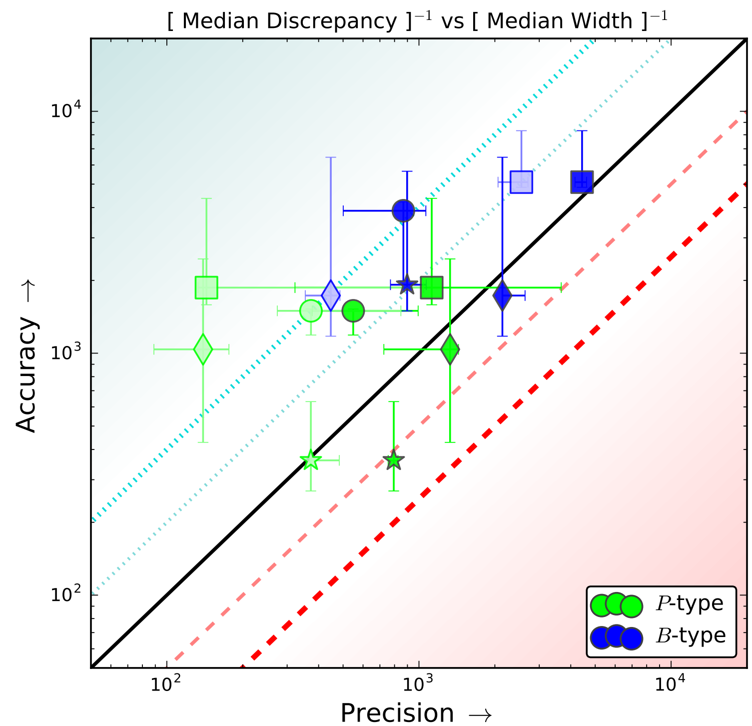


Figure 3.13. Z-scores, accuracy, and precision of fitted eclipse depths when we use sensitivity variations on the pixel that are either P-like (i.e. Equation 3.3) or B-like (i.e. defined at the knot locations and interpolated) to create a light curve. The circles and squares are taken from Figure 3.10. For the case shown by the blue diamond, accounting for time-correlated residuals changes the accuracy and precision from near-ideal (solid black line) to very conservative (outer dotted line and cyan region)—using these β factors can be dicey. BLISS is the more precise model for light curves with B-like sensitivity variations, but the polynomial fits have more predictive power when these variations are used in our main case (stars).

We find that BLISS fits data with significant eclipses, and (some) light curves made from BLISS-like sensitivity variations, better than the second-order polynomial. Strangely though, the polynomial model is at least as precise as BLISS in many cases we test, and is even preferred in several of them. As for using β factors to inflate uncertainties, we get mixed results: these can change dubious fits into near-ideal ones (e.g. green star in Figure 3.10), but can also make unreliable fits look reasonable (blue square in Figure 3.12) and reliable fits far too conservative (e.g. blue diamond in Figure 3.13).

3.5. Discussion

3.5.1. Kernel Regression

We have focused on BLISS because it is easy to adapt the method to a full Jump-type MCMC. Many researchers use BLISS to model intra-pixel sensitivity variations in IRAC data, but there are other non-parametric methods as well. The original approach is Kernel Regression (KR), first used on the GJ 436 system by Ballard et al. (2010). To measure the transit depth at a known point in a long time-series, the out-of-transit data (i.e. a control) were used to model the detector once at the start of the analysis. This detector model was then used to correct the in-transit data.

Since then, KR has been applied to phase observations, where the signal spans the entire observed baseline and there are no control data (Knutson et al., 2012; Lewis et al., 2013). Researchers have therefore adopted an optimization strategy similar to BLISS: at every MCMC step, the observed flux is divided by the current astrophysical model and KR is applied to the residuals. The KR implemented by Knutson et al. (2012) and Lewis et al. (2013) also includes the width of the point-spread function, but this does not change the similarity between KR and BLISS. In fact, recent studies have used this width of the point-spread function in tandem with BLISS (Lanotte et al., 2014; Demory et al., 2016b,a).

KR differs superficially from BLISS because it has no obvious detector parameters, making it less clear how to adapt KR to full Jump-type fits. Nonetheless, one can estimate the effective number of parameters as suggested by Footnote 2 of Hansen et al. (2014), typically of order 10^2. Given the conceptual similarities between BLISS and KR, it is possible that our results about the former apply to the latter.

3.5.2. Precision of Polynomial Models

From the sets of fitted eclipse depths in Figures 3.10–3.13, it is surprising that a second-degree polynomial is as (or more) precise than BLISS many times. That does not tend to happen with real *Spitzer* data: in both Stevenson et al. (2012a) and Blecic et al. (2013), the BLISS models are more precise than any polynomials the authors test through order $n = 6$. In fact, the choice between the models seems so clear that many works in Table 3.1 do not mention polynomials at all. Remember, BLISS is more precise and often more accurate when we test light curves with significant eclipses (squares in Figure 3.10) or those made with BLISS-like sensitivity variations (Figure 3.13). But in some cases, one is better off modeling the sensitivity with a low-order polynomial—there are several thoughts about why this can happen.

It would be great to fit real IRAC light curves that have $\sim 10^5$ data with all of these sensitivity models, but as mentioned in Sections 3.4.2–3.4.3, this is not computationally feasible (more modest IRAC measurements could work, though). Instead we *mimic* these light curves and fits by using realistic parameters for $A(t)$ and $D(t)$, choosing a reasonable BLISS mesh for the synthetic data, and running our MCMC chains until we get many independent samples. But maybe having more data and BLISS knots simply is different, even though the sensitivity at the knot locations and the interpolated maps are mostly accurate in our tests (Section 3.4.3.1). If so, both parameters likely have to increase as we do not see BLISS improve when changing only the data (Figure 3.11). This is not because our bin sizes are too large, either. Our BLISS knots are spaced ~ 0.02 pixels apart in both x and y—in other studies this number ranges from smaller (e.g. 3.6 μm data in Diamond-Lowe et al., 2014) to larger (e.g. 5.8 μm data in Blecic et al., 2013).

Also, when we set the amplitude of the detector signal in Equation 3.4, we do not pick *when* the sensitivity will rise and fall—that would mean explicitly choosing the centroids. Instead, the pointing model (Ingalls et al., 2016) and sensitivity map (Appendix A.1) that we use determine how the detector signal looks. If this $D(t)$ is flat with a single large spike or dip at one moment, most data is uncorrupted by the pixel's sensitivity (we avoid these signals for MCMC fits; Section 3.4.3.4). Moreover, it is mostly chance that the $V(x, y)$ from Equation 3.3 is very featured near the centroids. It only happens with particular sets of coefficients, true for any high-order polynomial. This means the sensitivity variations under the centroids can look quadratic *even if* the detector signal looks complex. Either way, our second-order polynomial model could give a good fit. Generating polynomial $V(x, y)$ by selecting their roots, not coefficients, would thus be interesting test cases. When the pixel's sensitivity is BLISS-like none of this should matter, though these variations have other issues (Section 3.4.3.4).

One possibility comes straight from Stevenson et al. (2012a): the MCMC would not converge because the model for the exponential ramp had strong, non-linear correlations. To solve this, the authors orthogonalized the ramp parameters, then transformed back to the first model after the MCMC to get the uncertainties. We have no ramping in our synthetic light curves, but similar correlations could happen in our model for $A(t)$. We see evidence of this when we try fitting a DC offset in Equation 3.5, which is why we fix the mean of the phase function to unity. There are a couple of problems with this idea, though. Our MCMC chains have little trouble stabilizing, even if `emcee` takes a long time to get there. Also, we fit the astrophysical signal identically for all three of our sensitivity models. Even if our parameters are not ideal, each fit should be affected the same way by $A(t)$, meaning this is probably not why the polynomial models are more precise.

Above all else, one might say we simply have not tested the “right” or enough types of light curves. At worst this means we explored some parts of parameter space that differ from real observations. But again, we have chosen a variety of cases based on real IRAC data, fit five or ten examples of each case to have statistics, and even try sensitivity variations more suited to BLISS (Figure 3.13). Our results in Section 3.4.3.4 are also only about a third of all our trials; we tested other combinations of parameters. Indeed, BLISS can be more precise than a second-order polynomial, such as by having a significant eclipse, a very large detector amplitude, or BLISS-like variations in sensitivity on the pixel. Yet polynomials are sometimes preferred and the BLISS method is fundamentally the same each time—we find this odd. As mentioned above, testing types of light curves that extend our cases would be a good way to see if any trends here continue in general.

3.5.3. Modeling IRAC Noise

One might ask what Section 3.4.3 means for dealing with detector signals in *Spitzer* data. A potential view is to only use non-parametric methods that properly marginalize over the detector behavior, like ICA (Waldmann, 2012) and Gaussian Processes (Gibson et al., 2012), or rather use viable parametric methods such as PLD (Deming et al., 2015; Ingalls et al., 2016). We find a polynomial model is often as precise as BLISS at fitting our synthetic eclipses. For real light curves, one could use high-order polynomials for the sensitivity (e.g. $n \geq 7$) and fit every term directly. The number of parameters would be similar to some of our Jump-BLISS models—we find `emcee` can handle this many jump dimensions.

However, we also have multiple cases (e.g. Figures 3.10 and 3.13) that match other studies where BLISS is the more precise choice. Non-parametric models can misfit uncertainties or bias a result (Figure 3.1), which is disconcerting because one cannot know how accurately BLISS fits real data. We find, though, that BLISS tends to be more accurate than precise (i.e. conservative) at fitting eclipse depths. This result could hold when the routine uses significantly more data and knots (Figure 3.11). And while sources of red noise can ruin a fit, a small leftover amount in a light curve *could* be beneficial for the predictive power of BLISS (Figure 3.12).

Yet there are other problems. While β factors (e.g. Pont et al., 2006; Cowan et al., 2012b) are an expedient way to account for time-correlated residuals, these can turn a reasonable uncertainty on an eclipse depth into overly conservative. BLISS does not predict the centering precision of an observation, though we hypothesize the routine may be able to indicate the length scale of the sensitivity variations. Methods to properly size the BLISS mesh, in this paper and Stevenson et al. (2012a), also may not work as intended (Sections 3.4.3.1 and 3.4.3.4). Furthermore, BLISS is often used with the Photometry for Orbits, Eclipses, and Transits pipeline (POET; Stevenson et al., 2012a; Cubillos et al., 2013), as in Ingalls et al. (2016). This proprietary code in part reduces pixelation of the detector by using flux-conserving, interpolated photometry (e.g. Figures 2 and 5 of Stevenson et al., 2012a). But it is unclear if this influences the apparent sensitivity variations on the pixel, and so makes BLISS more necessary to correct for them. Luckily, the large variations at 3.6 μm with IRAC should mean the influence of pixelation, or POET, is more negligible in this channel (e.g. Figure 7 of Blecic et al., 2013).

In any case, if a light curve has distinct astrophysical and detector signals, then one could likely use many approaches to model $D(t)$ reasonably. In contrast, a gradual rise and fall in detector sensitivity while observing a planet could be confused with phase variations. As Ingalls et al. (2016) shows, multiple sensitivity models can all fit the same eclipse depth (of XO-3b) well. We expect and see that this sometimes happens with our synthetic light curves.

There probably is no ideal method for handling the sensitivity variations in *Spitzer* IRAC data, and BLISS has both positive and negative qualities. We deem that the good significantly outweighs the bad, though, and suggest that using BLISS as a shortcut can be a practical approach.

3.6. Conclusions

We have performed MCMC fits on synthetic eclipse data to test how accurate and precise BLISS mapping is for modeling intra-pixel sensitivity variations in *Spitzer* IRAC light curves. BLISS mapping is a non-parametric method, meaning it uses no jump parameters during the MCMC to model the detector signal. This is an expedient approximation that is not statistically sound in principle. Nonetheless, BLISS mapping has been widely used without rigorous testing on synthetic data.

Optimizing nuisance parameters, instead of marginalizing over them, can give both imprecise and inaccurate estimates for other parameters of interest. Even in our toy example with simple posteriors, we find that fitted uncertainties can still be too small, by a factor of 2. In BLISS mapping, the estimated sensitivities at the knots—and so the interpolated maps—become inaccurate for the data when the photon noise is low. The maps also start fitting noise when the average data per knot is ~ 10 or less. However, in many reasonable cases, the knot values match the intrinsic sensitivity to within the photon noise and the maps give good fits to the detector signal.

Furthermore, standard BLISS mapping is a viable shortcut to the rigorously Bayesian Jump-BLISS mapping. Both methods return similar estimates for the astrophysical model, and the knots in standard BLISS mapping behave like actual jump parameters. Curiously, our low-order polynomial model is often as precise as BLISS mapping at fitting eclipse depths, yet the latter is preferred for high-significance eclipses and more featured sensitivity variations. We also find that using the β method to inflate uncertainties does not always increase the predictive power of fits.

In our tests, BLISS mapping does not predict the centering precision of a data set. Selecting a proper number of knots can require fine-tuning—proposed methods may not work in general. But, we find that BLISS mapping usually fits eclipse depths more accurately than precisely (i.e. conservatively), a potential benefit against low levels of extra red noise in the light curve. Overall, therefore, BLISS mapping can be an acceptable way to model *Spitzer* IRAC sensitivity variations.

3.7. Interlude II

Non-parametric methods have been a part of statistics for many decades; it is no surprise they have worked their way into astronomical data analysis. Whatever one's stance on their use, a few things are reasonably clear from our foray into pixel sensitivities. Optimizing nuisance parts of a model works very well—except when it fails. After all, it can be quite easy to misfit pieces of light curves that we care about. That is *with* the benefit here of knowing those particular models are poor, a luxury that simply does not happen very often in the real world. But does this mean all non-parametric ways to clean up light curves are evil? Of course not: in practice BLISS seems alright. It just means these methods have to be used carefully, much like using bleach as a sanitizer. When handled properly, any sensitivity analysis gives us more confidence about the science that stems from those published results. The outlook is bright, allowing us to take a culminating step into empirical trends from transiting planet observations. Describing terrestrial worlds is the goal, but hot Jupiters have a relatively large amount of eclipse and phase data available for now, which we wield to quantitatively characterize their atmospheres.

CHAPTER 4

Balancing the Energy Budget of Short-Period Giant Planets: Evidence for Reflective Clouds and Optical Absorbers

This chapter is adapted from Schwartz, J. C., and Cowan, N. B. 2015, MNRAS, 449, 4192.

The section on updates is partially adapted from Schwartz, J. C., Kashner, Z., and Cowan, N. B., In Prep.

4.1. Introduction

Mature planets on short-period orbits have energy budgets dominated by incoming radiation rather than internal heat. Their atmospheric temperatures are therefore a function of both the absorption of incident stellar energy and its transport before re-emission into space.

Absorbed energy is solely a matter of incident stellar flux and the planet's Bond albedo, A_B. Ironically, it is difficult to constrain Bond albedo (the fraction of stellar energy that is reflected) through observations of reflected light. In order to convert an optical geometric albedo (visible light reflected towards the illuminating star) into a Bond albedo, one must make assumptions about a planet's reflectance spectrum, scattering phase function, and spatial inhomogeneity (Hanel, 2003). Bond albedo is more readily obtained from thermal measurements via energy balance.

Heat transport is more complicated, but tends to move energy from hot to cold: vertically upwards, from equator to pole (for planets with small obliquity), from summer hemisphere to winter hemisphere (for planets with non-zero obliquity), and from day to night (for planets with slow rotation). Due to strong tides, short-period planets are expected to have zero obliquity and slow rotation, and most are on circular orbits. Hot Jupiters on circular orbits are further expected to be tidally-locked, with one side permanently facing the host star and the other forever dark. As such, the atmosphere tends to transport heat from the dayside to the nightside, and from equatorial regions to the poles (for a recent review of hot Jupiter atmospheric dynamics, see Heng & Showman, 2014).

The atmospheric temperature of a planet is generally a function of four variables: longitude, latitude, pressure (or height), and time. The time-dependence can usually be neglected for hot Jupiters because they exhibit minimal weather (Agol et al., 2010; Knutson et al., 2011; Wong et al., 2014). Moreover, hot Jupiters on circular orbits are expected to have a 3D fixed temperature structure with respect to the sub-stellar location, regardless of whether they are tidally locked (Rauscher & Kempton, 2014; Showman et al., 2014). This motivates using a star-based coordinate system with the prime meridian facing the star, and allows us to use orbital phase as a proxy for longitude (Cowan & Agol, 2008). The latitudinal temperature-dependence of a hot Jupiter is inaccessible unless one can measure higher-order phase modulation (Cowan et al., 2013) or utilize occultation mapping (Majeau et al., 2012; de Wit et al., 2012). Finally, the vertical temperature structure is in principle accessible via emission spectroscopy: wavelengths at which the atmosphere is relatively transparent will probe deeper layers, and vice versa.

Multi-wavelength thermal phase variations of a hot Jupiter on a circular orbit therefore amount to brightness temperature measurements as a function of orbital phase and wavelength (e.g., Stevenson et al., 2014c). If one is solely interested in the global properties of the planet—namely Bond albedo and day-to-night heat transport—then one can further simplify the problem by combining brightness temperatures at each orbital phase to obtain a bolometric flux, and hence an effective temperature *at that phase*. Note that Solar System planets tend to have effective temperatures that are either uniform from any vantage point, or which vary based on the latitude of the observer due to imperfect poleward heat transport. For short-period exoplanets, on the other hand, the principal temperature gradient is between day and night, and dayside effective temperatures are often hundreds to thousands of degrees greater than their nightside counterparts. The final simplification we make is therefore to treat the planet as two horizontally isothermal hemispheres: a dayside and a nightside. The effective temperatures of each hemisphere are simply the weighted mean of the atmospheric temperatures on that side of the planet.

4.1.1. Previous Work

Cowan & Agol (2011b) used broadband infrared eclipse measurements of 24 hot Jupiters to demonstrate that they have generally low Bond albedos ($A_B < 0.5$), and that the hottest planets have extremely low albedos and/or poor day–night heat transport efficiency, ε.

It is possible to break the albedo-transport degeneracy by combining dayside thermal constraints with measurements of either nightside thermal emission or dayside reflected light. Cowan et al. (2007) used an 8.0 μm eclipse depth and phase amplitude from the Spitzer Space Telescope, combined with an optical eclipse measurement from the MOST satellite, to constrain the energy budget of HD 209458b; they inferred a small Bond albedo (absorption of almost all light that shines on it) and a high day-night heat transport efficiency (nightside not much cooler than the dayside).

Cowan et al. (2012c) used 8 infrared dayside and 2 mid-infrared nightside measurements to constrain the albedo and recirculation of WASP-12b; they found the planet has a modest Bond albedo (~ 0.25) and low heat transport efficiency ($\lesssim 0.1$).

Stevenson et al. (2014c) used phase-resolved emission spectroscopy taken with WFC3 from the Hubble Space Telescope to map the atmospheric thermal structure of WASP-43b, finding low Bond albedo (0.06–0.25) and no heat redistribution (for recent reviews of exoplanet atmospheric observations, please see Burrows, 2014; Bailey, 2014).

Our work is organized as follows: in Section 4.2 we use published eclipse depths at multiple infrared wavelengths to infer effective dayside temperatures for fifty planets, with more than twice the data as Cowan & Agol (2011b). In Section 4.3 we consider the subset of planets for which we can break the albedo-recirculation degeneracy. We first tackle the six planets with thermal measurements of both eclipses and phase variations (Section 4.3.1), then the nine planets for which reflected light measurements are available in addition to dayside thermal constraints (Section 4.3.2). We discuss our results in Section 4.4.

4.2. Dayside Energy Budget

Inferring the effective dayside temperature of a planet requires combining eclipse depths at thermal wavelengths, which we define as those longward of 0.8 μm. We have updated the sample from Hansen et al. (2014), now including fifty planets with a minimum

Table 4.1. Short-period giant planets with a minimum of two published eclipse observations at infrared wavelengths (non-detections are not included).

Infrared Multi-Eclipse Planets	
CoRoT-1b	TrES-2b
CoRoT-2b	TrES-3b
GJ 436b	TrES-4b
HAT-P-1b	WASP-1b
HAT-P-2b	WASP-2b
HAT-P-3b	WASP-3b
HAT-P-4b	WASP-4b
HAT-P-6b	WASP-5b
HAT-P-7b	WASP-6b
HAT-P-8b	WASP-8b
HAT-P-12b	WASP-12b
HAT-P-19b	WASP-14b
HAT-P-20b	WASP-17b
HAT-P-23b	WASP-18b
HAT-P-32b	WASP-19b
HD 149026b	WASP-24b
HD 189733b	WASP-26b
HD 209458b	WASP-33b
KELT-1b	WASP-39b
Kepler-5b	WASP-43b
Kepler-6b	WASP-48b
Kepler-12b	XO-1b
Kepler-13Ab	XO-2b
Kepler-17b	XO-3b
TrES-1b	XO-4b

of two published infrared eclipse measurements (additions include HAT-P-19b, HAT-P-20b, HAT-P-32b, WASP-6b, WASP-26b, WASP-39b; Mahtani et al., 2013; Deming et al., 2015; Zhao et al., 2014, Kammer et al. in prep.). Our data predominantly consist of broadband photometry, but we also include spectroscopic emission measurements when they are at complementary wavelengths (e.g. Ranjan et al., 2014; Wilkins et al., 2014; Crouzet et al., 2014; Stevenson et al., 2014c). The planets from our sample are listed in Table 4.1.

4.2.1. Brightness Temperatures

Thermal emission at different wavelengths originates from different layers in the planet's atmosphere, which have different temperatures. One can define a brightness temperature at each observed wavelength, $T_b(\lambda)$: this is the temperature that a blackbody must have in order to emit at the same intensity as the planet.

Inverting the Planck function, we obtain the following expression for brightness temperature (Cowan & Agol, 2011b):

$$T_b(\lambda) = \frac{hc}{\lambda k}\left[\log\left(1 + \frac{e^{hc/\lambda k T_*} - 1}{\psi(\lambda)}\right)\right]^{-1}, \tag{4.1}$$

where $\psi(\lambda)$ is the relative intensity of the planet to that of its host star and T_* is the stellar effective temperature, meaning we treat the star as a blackbody. For dayside measurements, $\psi(\lambda)$ is the ratio of eclipse depth to transit depth, $\delta_{\mathrm{ecl}}/\delta_{\mathrm{tr}}$, while for nightside measurements it is the ratio of nigthside flux to transit depth, $(\delta_{\mathrm{ecl}} - \delta_{\mathrm{var}})/\delta_{\mathrm{tr}}$, where δ_{var} is the phase variation amplitude. Published data therefore allow us to compute dayside (and, when appropriate, nightside) brightness temperatures for each waveband in which a planet has been observed.

4.2.2. Aggregate Emission Spectrum

The broadband emission spectra of most individual planets are consistent with isothermal atmospheres (Hansen et al., 2014). It is possible, however, to construct an aggregate emission spectrum of all fifty planets in the hopes of revealing molecular absorption features too faint to detect in any individual planet's spectrum.

If some planets have temperature inversions while others do not, this type of averaging could actually wash out molecular signatures, which would appear in absorption for some planets and emission for others. However, the first and most statistically significant case of a hot Jupiter temperature inversion (Knutson et al., 2008) has not been borne out by new measurements nor reanalysis of the originals (Zellem et al., 2014; Diamond-Lowe et al., 2014; Schwarz et al., 2015). Moreover, a systematic study of *Spitzer* eclipse measurements found that they have not been as accurate as advertised (Hansen et al., 2014), suggesting the temperature inversions reported in most hot Jupiter atmospheres may simply be due to confirmation bias.

The aggregate emission spectrum for the fifty hot Jupiters is shown in Figure 4.1. We normalize the brightness temperature spectrum of each planet in our sample, then determine the median and uncertainty on the mean at each wavelength. This "stacking" is only useful, however, at wavelengths for which there are observations of many planets (currently 1.15, 1.65, 2.25, 3.6, 4.5, 5.8, and 8.0 microns).

There are no significant features in the average spectrum, not even the trend towards higher brightness temperatures at shorter wavelengths reported by Burrows (2014). It is worth noting that Figure 4 of Burrows (2014) used data from fewer planets, and was normalized differently: the *equilibrium temperature* of each planet was divided out, rather than its actual *dayside effective temperature*. Recall that a planet's equilibrium temperature is what one would expect for a planet with zero Bond albedo and uniform temperature; it is merely a convenient theoretical quantity proportional to the irradiation temperature, T_0, that we utilize in this work. The dayside effective temperature, on the

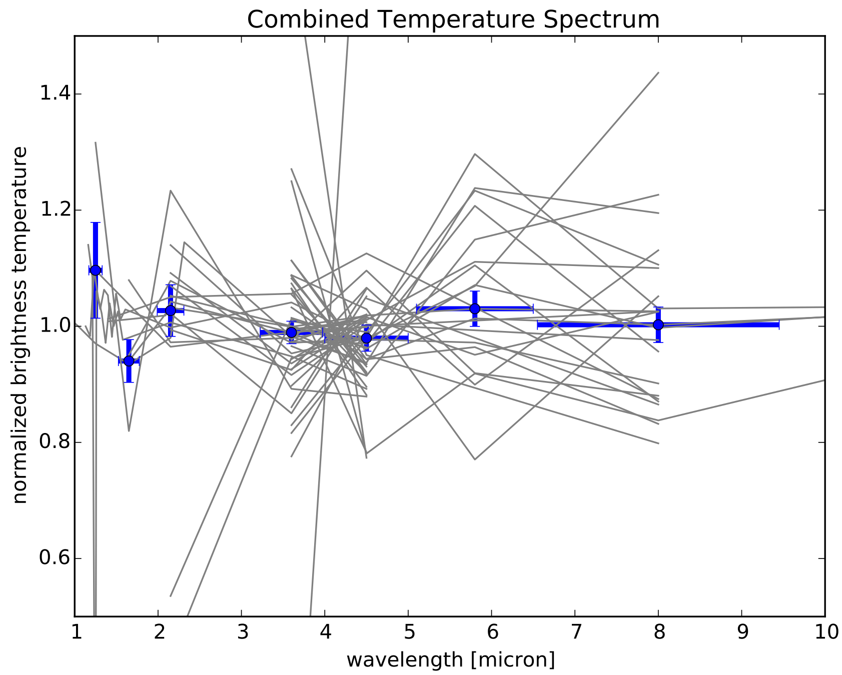


Figure 4.1. Average broadband emission spectrum for fifty short-period giant planets (blue), plotted with the emission spectra of individual planets (gray). The spectrum of each planet is normalized to its mean brightness temperature. The aggregate spectrum is the median normalized brightness temperature at each wavelength, where the uncertainty bars denote the uncertainty on the mean (as opposed to the standard variation of the spectra at that wavelength.)

other hand, is the weighted mean brightness temperature of the planet's dayside, as described in Section 4.2.3. Most hot Jupiters have dayside effective temperatures greater than their equilibrium temperature (dotted line in Figure 4.2) due to imperfect day–night heat transport. As a result, the dayside of a hot Jupiter emits somewhat more in the mid-IR—and considerably more in the near-IR—than one would predict based on its equilibrium temperature. In any case, we agree with Burrows (2014) that there are no signs of molecular absorption features in the aggregate spectrum. Since molecules are

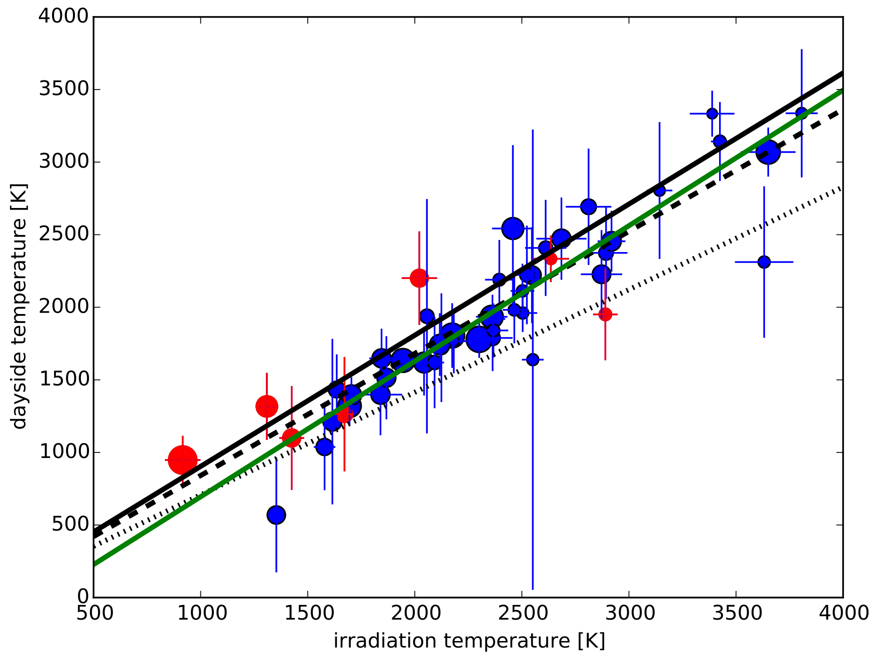


Figure 4.2. Dayside effective temperature versus irradiation temperature for all giant planets with multiple dayside infrared eclipses, estimated by Monte Carlo using hybrid EWM-PWM calculation and inflating observational uncertainties by $f_{\rm sys} = 3$ where applicable (Hansen et al., 2014). Uncertainty bars for both temperatures are shown, while dot size is proportional to the fraction of expected planetary emission that falls within observed wavebands. Red symbols denote eccentric planets ($e > 0.1$). The solid, dashed, and dotted lines correspond respectively to maximum dayside temperature, $T_d = (2/3)^{1/4}T_0$, uniform dayside with zero nightside temperature, $T_d = (1/2)^{1/4}T_0$, and equilibrium temperature, $T_d = (1/4)^{1/4}T_0$. The trend line is shown in green; it suggests that hotter planets have disproportionately hot daysides.

undoubtedly present in the atmospheres of exoplanets, we conclude that their absorption features are being muted by vertically isothermal atmospheres, optically thick cloud, or both.

4.2.3. Effective Temperatures

While brightness temperatures of brown dwarfs are strongly wavelength-dependent in the near-infrared (Faherty et al., 2014; Biller et al., 2013), the external, asymmetric heating experienced by hot Jupiters produces dayside atmospheres that are relatively isothermal in the vertical direction (Fortney et al., 2006). This results in relatively featureless dayside emission spectra, which are amenable to model-independent estimates of bolometric flux and hence effective dayside temperature. If the nightsides of hot Jupiters have greater vertical temperature structure, then nightside effective temperature estimates are less reliable.

There is no universal way to derive the effective temperature of a planet from a collection of brightness temperatures. We therefore consider two methods with different physical motivations. In the first method, each brightness temperature is weighted by the inverse square of its respective uncertainty ($\omega_i = 1/\sigma_i^2$), so eclipse depths with small relative uncertainties contribute more to the inferred effective temperature. We call this the error-weighted mean (EWM) effective temperature. This method has the advantage of being robust to occasional outlier eclipse depths, but implicitly assumes that short-period planets have Planck-like broadband spectra.

The second method, which we only apply to dayside measurements, weighs the brightness temperatures by the expected integrated power in that waveband: $\omega_i = P_i = \int_{\lambda_1}^{\lambda_2} B(T_{\rm est}, \lambda) d\lambda$, where $B(T_{\rm est}, \lambda)$ is the Planck function for the planet's estimated dayside effective temperature. To bypass an iterative solution, we adopt $T_{\rm est} = (\frac{1}{2})^{1/4} T_0$, where T_0 is the irradiation temperature: $T_0 \equiv T_* \sqrt{R_*/a}$ where R_* is the stellar radius and a is the semi-major axis. This is an excellent match (the dashed line in Figure 4.2) to the actual dayside effective temperature of most planets in our sample. We call this the power-weighted mean (PWM); it is identical in spirit to the linear interpolation method of Cowan & Agol (2011b), but is easier to implement and runs faster. The PWM gives more weight to measurements near the peak of the planet's Planck function and should, in the limit of high spectral coverage, produce accurate effective temperatures even if planets have broadband spectral features.

Note that both EWM and PWM are biased in favor of broadband measurements: these observations tend to have smaller uncertainties, and they capture more of the planet's expected blackbody emission. The two methods produce generally consistent effective temperature estimates, which is a testament to the fact that most current dayside emission spectra are approximately Planck-like.

We use a 10^4-step Monte Carlo analysis to estimate uncertainties in dayside effective temperatures. At each step in the Monte Carlo, we randomly vary the stellar effective temperature, transit depth, eclipse depth, and scaled semi-major axis, $a_* \equiv a/R_*$, according to their uncertainties. We use the published uncertainties for all of the above, except for single-epoch broadband eclipse depths where we inflate the published uncertainty by the factor $f_{\rm sys} = 3$ (such measurements have historically been less accurate than advertised; Hansen et al., 2014). We also compute each planet's irradiation and brightness temperatures (following Equation 4.1). We then estimate each planet's dayside effective temperature; to hedge our bets, we use the EWM for half of the MC steps, and the PWM for the other half. The resulting relationship between T_0 and T_d is shown in Figure 4.2. The median property is plotted, and uncertainty bars denote standard deviation from the MC.

4.3. Global Energy Budget

Based on dayside effective temperatures alone, one cannot simultaneously specify Bond albedo and heat recirculation efficiency. This degeneracy can be broken by supplementing thermal eclipses with one of two measurement types: phase variations at infrared wavelengths, or eclipse depths at visible wavelengths. Table 4.2 lists the published data for the fifteen planets which fall into one or both of these categories: observed wavelength and bandwidth, eclipse depths, and phase amplitudes. Cyan-colored entries are measurements exempt from the $f_{\rm sys} = 3$ uncertainty inflation of Hansen et al. (2014).

Table 4.2. Eclipse depths and phase amplitudes for our restricted planetary samples where degeneracy between albedo and heat recirculation can be resolved. Observations are denoted by central wavelength and bandwidth; measurements in cyan are exempt from the uncertainty inflation $f_{\rm sys} = 3$ of Hansen et al. (2014).

Planet	Wavelength (μm)	Eclipse Depth		Phase Amplitude	
CoRoT-1b	0.60(42)	1.6(6)	$\times 10^{-4}$		
	1.65(25)	1.45(49)	$\times 10^{-3}$		
	2.10(2)	2.8(5)	$\times 10^{-3}$		
	2.15(32)	3.190(405)	$\times 10^{-3}$		
	3.60(75)	4.15(42)	$\times 10^{-3}$		
	4.5(10)	4.82(42)	$\times 10^{-3}$		
CoRoT-2b	0.60(42)	6(2)	$\times 10^{-5}$		
	1.4(6)	3.95(57)	$\times 10^{-4}$		
	1.65(25)	8.50(283)	$\times 10^{-4}$		
	2.15(32)	1.6(9)	$\times 10^{-3}$		
	3.60(75)	3.55(20)	$\times 10^{-3}$		
	4.5(10)	4.75(19)	$\times 10^{-3}$		
	8.0(29)	4.09(80)	$\times 10^{-3}$		
HAT-P-7b	0.65(40)	7.12(15)	$\times 10^{-5}$	7.33(27)	$\times 10^{-5}$
	3.60(75)	9.8(17)	$\times 10^{-4}$		
	4.5(10)	1.59(22)	$\times 10^{-3}$		
	5.8(14)	2.45(31)	$\times 10^{-3}$		
	8.0(29)	2.25(52)	$\times 10^{-3}$		

Planet	Wavelength (μm)	Eclipse Depth		Phase Amplitude	
HD 149026b	3.60(75)	4.0(3)	$\times 10^{-4}$		
	4.5(10)	3.4(6)	$\times 10^{-4}$	2.23(58)	$\times 10^{-4}$
	5.8(14)	4.4(10)	$\times 10^{-4}$		
	8.0(29)	3.7(8)	$\times 10^{-4}$	2.3(7)	$\times 10^{-4}$
	16(5)	8.5(32)	$\times 10^{-4}$		
HD 189733b	0.37(16)	1.26(37)	$\times 10^{-4}$		
	0.51(12)	1(34)	$\times 10^{-6}$		
	1.4(6)	9.6(39)	$\times 10^{-5}$		
	2.15(32)	2(2)	$\times 10^{-4}$		
	3.60(75)	1.47(4)	$\times 10^{-3}$	1.240(61)	$\times 10^{-3}$
	4.5(10)	1.790(38)	$\times 10^{-3}$	9.82(89)	$\times 10^{-4}$
	5.8(14)	3.10(34)	$\times 10^{-3}$		
	6.45(210)	2.200(62)	$\times 10^{-3}$		
	8.0(29)	3.44(36)	$\times 10^{-3}$		
	10.5(60)	3.560(67)	$\times 10^{-3}$		
	16(5)	5.51(30)	$\times 10^{-3}$		
	24(9)	5.36(27)	$\times 10^{-3}$	1.3(3)	$\times 10^{-3}$
HD 209458b	0.5(3)	7(9)	$\times 10^{-6}$		
	2.15(32)	1.5(15)	$\times 10^{-4}$		
	3.60(75)	9.4(9)	$\times 10^{-4}$		
	4.5(10)	1.3900(705)	$\times 10^{-3}$	1.090(115)	$\times 10^{-3}$
	5.8(14)	3.01(43)	$\times 10^{-3}$		
	8.0(29)	2.40(26)	$\times 10^{-3}$	7.50(375)	$\times 10^{-4}$
	24(9)	3.38(26)	$\times 10^{-3}$		
Kepler-5b	0.65(40)	1.86(36)	$\times 10^{-5}$	1.93(58)	$\times 10^{-5}$
	3.60(75)	1.03(17)	$\times 10^{-3}$		
	4.5(10)	1.07(15)	$\times 10^{-3}$		
Kepler-6b	0.65(40)	1.11(40)	$\times 10^{-5}$	1.72(43)	$\times 10^{-5}$
	3.60(75)	6.9(27)	$\times 10^{-4}$		
	4.5(10)	1.51(19)	$\times 10^{-3}$		
Kepler-7b	0.65(40)	3.870(835)	$\times 10^{-5}$	4.8(13)	$\times 10^{-5}$
	3.60(75)	3.08(103)	$\times 10^{-4}$		
	4.5(10)	5.05(168)	$\times 10^{-4}$		
Kepler-13Ab	0.65(40)	1.720(18)	$\times 10^{-4}$	1.5200(105)	$\times 10^{-4}$
	2.15(32)	1.22(51)	$\times 10^{-3}$		
	3.60(75)	1.56(31)	$\times 10^{-3}$		
	4.5(10)	2.22(23)	$\times 10^{-3}$		
TrES-2b	0.65(40)	7.7(18)	$\times 10^{-6}$	4.10(105)	$\times 10^{-6}$
	2.15(32)	6.2(12)	$\times 10^{-4}$		

Planet	Wavelength (μm)	Eclipse Depth		Phase Amplitude	
	3.60(75)	1.27(21)	$\times 10^{-3}$		
	4.5(10)	2.30(24)	$\times 10^{-3}$		
	5.8(14)	1.99(54)	$\times 10^{-3}$		
	8.0(29)	3.59(60)	$\times 10^{-3}$		
WASP-12b	0.90(15)	1.360(136)	$\times 10^{-3}$		
	1.04(12)	1.09(14)	$\times 10^{-3}$		
	1.25(16)	1.39(30)	$\times 10^{-3}$		
	1.38(55)	1.580(39)	$\times 10^{-3}$		
	1.4(6)	1.740(17)	$\times 10^{-3}$		
	1.65(25)	1.91(20)	$\times 10^{-3}$		
	2.15(32)	2.96(14)	$\times 10^{-3}$		
	2.220(34)	3.01(46)	$\times 10^{-3}$		
	2.320(27)	4.5(6)	$\times 10^{-3}$		
	3.60(75)	4.19(44)	$\times 10^{-3}$	3.20(33)	$\times 10^{-3}$
	4.5(10)	4.29(33)	$\times 10^{-3}$	3.92(16)	$\times 10^{-3}$
	5.8(14)	6.96(60)	$\times 10^{-3}$		
	8.0(29)	6.96(96)	$\times 10^{-3}$		
WASP-18b	3.60(75)	3.04(26)	$\times 10^{-3}$	2.96(11)	$\times 10^{-3}$
	4.5(10)	3.79(21)	$\times 10^{-3}$	3.66(9)	$\times 10^{-3}$
	5.8(14)	3.7(3)	$\times 10^{-3}$		
	8.0(29)	4.1(2)	$\times 10^{-3}$		
WASP-19b	0.685(530)	3.9(19)	$\times 10^{-4}$		
	0.91(20)	8.0(29)	$\times 10^{-4}$		
	1.29(8)	8.3(39)	$\times 10^{-4}$		
	1.6(4)	1.86(14)	$\times 10^{-3}$		
	1.65(25)	2.76(44)	$\times 10^{-3}$		
	2.15(32)	2.87(20)	$\times 10^{-3}$		
	3.60(75)	4.83(25)	$\times 10^{-3}$		
	4.5(10)	5.72(30)	$\times 10^{-3}$		
	5.8(14)	6.5(11)	$\times 10^{-3}$		
	8.0(29)	7.3(12)	$\times 10^{-3}$		
WASP-43b	1.4(6)	4.61(5)	$\times 10^{-4}$	4.68(4)	$\times 10^{-4}$
	1.65(25)	1.03(17)	$\times 10^{-3}$		
	2.15(32)	1.81(27)	$\times 10^{-3}$		
	3.60(75)	3.47(13)	$\times 10^{-3}$		
	4.5(10)	3.82(15)	$\times 10^{-3}$		

4.3.1. Full-Orbit Thermal Measurement

The first way to resolve the degeneracy between Bond albedo and heat recirculation is by combining thermal eclipse and phase measurements to infer the planet's nightside effective temperature. This requires phase variations at thermal wavelengths, again defined as longward of 0.8 μm. Such phase observations are more time-intensive than eclipses, and therefore less widely available. There are six planets with at least one published phase measurement: HD 149026b, HD 189733b, HD 209458b, WASP-12b, WASP-18b, and WASP-43b. As in Section 4.2, we include band-integrated spectroscopy when it complements photometric observations (only one case at the moment: Stevenson et al., 2014c). For non-detections, an $n\sigma$ upper limit of α is assumed to have a value and uncertainty of $\alpha/2$ and $\alpha/(2n)$ respectively. Table 4.2 shows the data for this sample.

Observational references are as follows: HD 149026b: Stevenson et al. (2012a); Knutson et al. (2009b); HD 189733b: Evans et al. (2013); Crouzet et al. (2014); Barnes et al. (2007); Knutson et al. (2012); Charbonneau et al. (2008); Todorov et al. (2014); Agol et al. (2010); Deming et al. (2006); Knutson et al. (2009c); HD 209458b: Rowe et al. (2008); Richardson et al. (2003); Knutson et al. (2008); Zellem et al. (2014); Cowan et al. (2007); WASP-12b: Föhring et al. (2013); Croll et al. (2014); Crossfield et al. (2012a); Stevenson et al. (2014a); Swain et al. (2013); Cowan et al. (2012c); WASP-18b: Maxted et al. (2013); Nymeyer et al. (2011); WASP-43b: Stevenson et al. (2014c); Wang et al. (2013); Zhou et al. (2014); Blecic et al. (2014).

4.3.1.1. Reflected Infrared Light. The light emanating from a planet's dayside is a combination of thermal emission and reflected starlight. We plot an example for CoRoT-2b at 1.4 μm in the bottom panel of Figure 4.3, assuming the geometric and Bond albedos

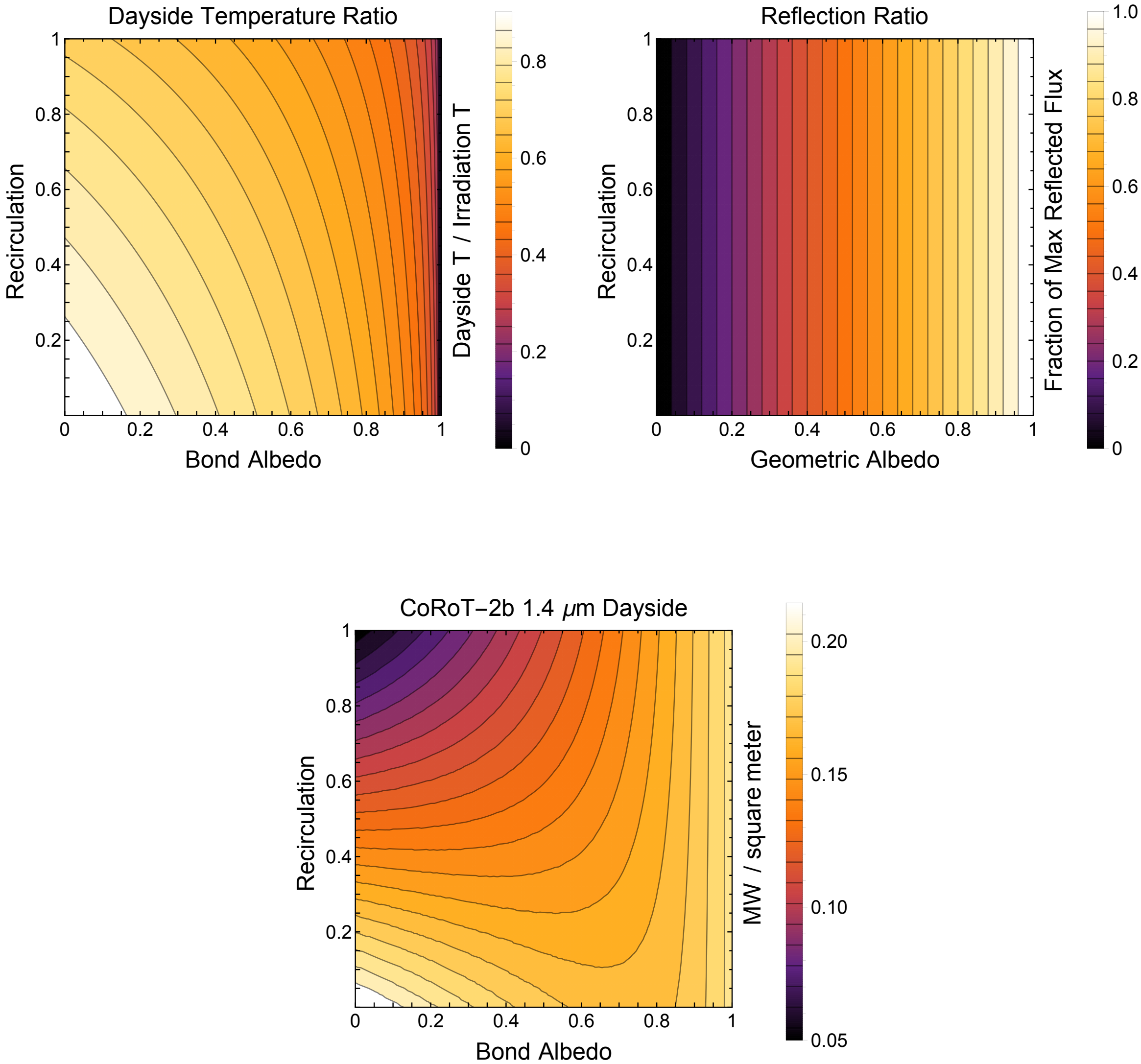


Figure 4.3. Dayside temperature (top left), reflection (top right), and dayside flux for CoRoT-2b at 1.4 μm (bottom), shown as a function of albedo and recirculation (assuming the geometric and Bond albedos are equal to one another; see Section 4.4.5 for caveats.) At high albedo, the NIR dayside flux roughly parallels reflected starlight, while at low albedo, the dayside flux is mostly thermal emission and hence depends on day-night heat transport.

are equal (though as described in Section 4.4.5 the proper conversion is more involved.) If the planet has low albedo in this scenario, then the 1.4 μm flux is almost entirely thermal emission and depends primarily on day–night heat transport. In the high albedo limit, on the other hand, the NIR flux is primarily reflected light and so varies linearly with the geometric albedo. In other words, even eclipse measurements at wavelengths *greater* than 0.8 μm are potentially contaminated by reflected starlight.

Furthermore, Figure 4.4 shows the reflected light contribution to dayside flux as a function of wavelength for a hypothetical gray planet (with temperature limits derived from Equation 4.2; see Section 4.3.1.2.) Reflected light dominates at ultraviolet wavelengths as expected, but its prevalence continues well through the near-infrared. For reasonable system parameters, reflected light contributes $\gtrsim 10$ per cent of the NIR flux (this reflected contribution goes up if T_* is increased, or if T_d or a_* are decreased.) Even if molecular absorption depresses the reflectance in certain bands, it is likely that eclipse measurements in NIR water opacity windows (J, H, and K) are contaminated by reflected light at the $\gtrsim 10$ per cent level.

4.3.1.2. Confidence Regions. In estimating effective dayside and nightside temperatures for each planet, we use Monte Carlo simulations with 5000 steps to propagate uncertainties. We assume uncertainties on observed quantities to be Gaussian and symmetrical; when asymmetrical uncertainty bars are reported, we adopt their mean. The planetary irradiation temperature is first computed as described in Section 4.2.3. Brightness temperatures are calculated for each appropriate measurement from Table 4.2 using Equation 4.1, propagating uncertainty on stellar effective temperature, eclipse depth, phase amplitude, and transit depth. For observations *not* listed in cyan (isolated eclipses,

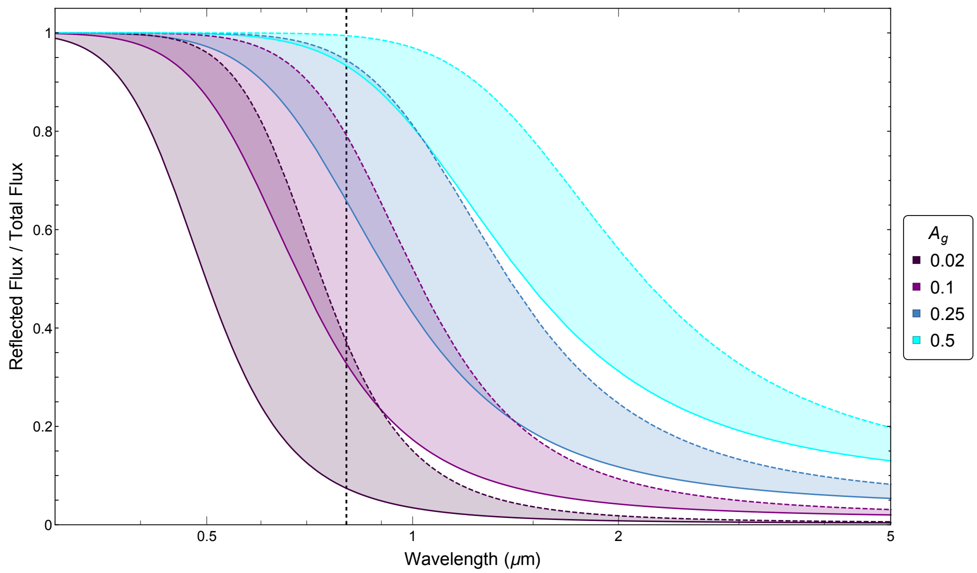


Figure 4.4. Contribution of planetary reflected light as a function of wavelength for different geometric albedos (lighter = higher), assuming blackbody emission and $q = \frac{5}{4}$ (see Section 4.4.5). This example planetary system assumes $T_* = 6100\ K$ and $a_* = 4.8$, both weighted means of the fifteen planets in Table 4.2. Albedo regions are bounded by dayside temperature limits using Equation 4.2: solid lines denote no recirculation ($\varepsilon = 0$), dashed lines denote perfect recirculation ($\varepsilon = 1$). The vertical dashed line indicates our cutoff of 0.8 μm between reflected light and thermal emission.

or partial phase curves), we conservatively inflate the published uncertainty by the factor $f_{\rm sys} = 3$ (Hansen et al., 2014). Reflected light contributions are subtracted from all dayside eclipse depths—using planetary radius, R_p, and semi-major axis—assuming infrared geometric albedos to be normally distributed with mean 0.07 and width 0.01 (this is the distribution of uncorrected optical geometric albedo values described in Section 4.4.5). In cases where a brightness temperature is calculated as 0 K for all MC steps, we assume 100 K uncertainty in subsequent propagations. We then compute the effective dayside and nightside temperatures using the EWM of the corresponding brightness temperatures

(as this requires no a priori temperature assumption and produces similar values to the PWM.)

Our parameterization of recirculation neglects any treatment of poleward heat transport. The most extreme meridional temperature profiles are either uniform in the North-South direction (perfect poleward transport) or $T \propto \cos^{\frac{1}{4}}\theta$, where θ is latitude (no poleward transport). The difference in effective temperature seen by an equatorial observer is $(1/4)^{1/4}\,T_0$ versus $(8/3\pi^2)^{1/4}\,T_0$, a 1 per cent discrepancy. We incorporate this worst-case systematic uncertainty in quadrature for all effective temperature estimates.

Understandably, the number of brightness temperature measurements at distinct wavelengths for a planet affects the accuracy of the effective temperature estimate. In a Monte Carlo analysis using J.J. Fortney atmospheric models, Cowan & Agol (2011b) estimated systematic errors of 7.6 per cent in effective temperature when only a single observation was used (note that we only consider planets with at least two measurements), down to approximately 2.5 per cent for four or more measurements. We conservatively adopt a similar sliding scale of 8 per cent down to 3 per cent systematic uncertainty in effective temperature over the same observation number range, again added in quadrature.

Once we have dayside and nightside effective temperatures—and realistic uncertainties—for the six exoplanets, it is possible to infer each planet's Bond albedo and day–night heat transport efficiency using the parameterization of Cowan & Agol (2011b):

$$T_d = T_0(1 - A_B)^{1/4}\left(\frac{2}{3} - \frac{5}{12}\varepsilon\right)^{1/4}, \tag{4.2}$$

and

$$T_n = T_0(1 - A_B)^{1/4} \left(\frac{\varepsilon}{4}\right)^{1/4}, \tag{4.3}$$

where both A_B and ε can take values between zero and unity.

We create χ^2 surfaces for each planet based on our estimated dayside and nightside effective temperatures and using Equations 4.2 and 4.3. We calculate χ^2 on a 101×101 grid in A_B and ε, then interpolate the intermediate values. The 1σ, 2σ, and 3σ confidence intervals are defined as $\Delta\chi^2 = \{1, 4, 9\}$ respectively above the minimum, $\chi^2_{\rm min}$, where $\chi^2_{\rm min} \approx 0$ for most planets because we have two constraints and two model parameters. Generating the χ^2 surfaces involves numerical integrations of Planck functions, which can be computationally intensive. We therefore create a database of relevant integrals; with 10^4 grid points tested per effective temperature, this database decreases computational time by more than 95 per cent.

We plot the 1σ confidence intervals for the six exoplanets with full-orbit thermal measurements in Figure 4.5. Each planet is colored according to irradiation temperature, essentially the incident stellar flux. Since these planets have benefited from intensive observational campaigns, omitting the $f_{\rm sys} = 3$ uncertainty inflation produces nearly identical confidence intervals.

4.3.2. Geometric Albedo Measurement

The alternative approach to resolving the albedo versus heat-transport degeneracy of thermal eclipses is to also consider eclipse measurements at optical wavelengths. For our purposes, optical eclipses are those shortward of 0.8 μm; these observations allow us to

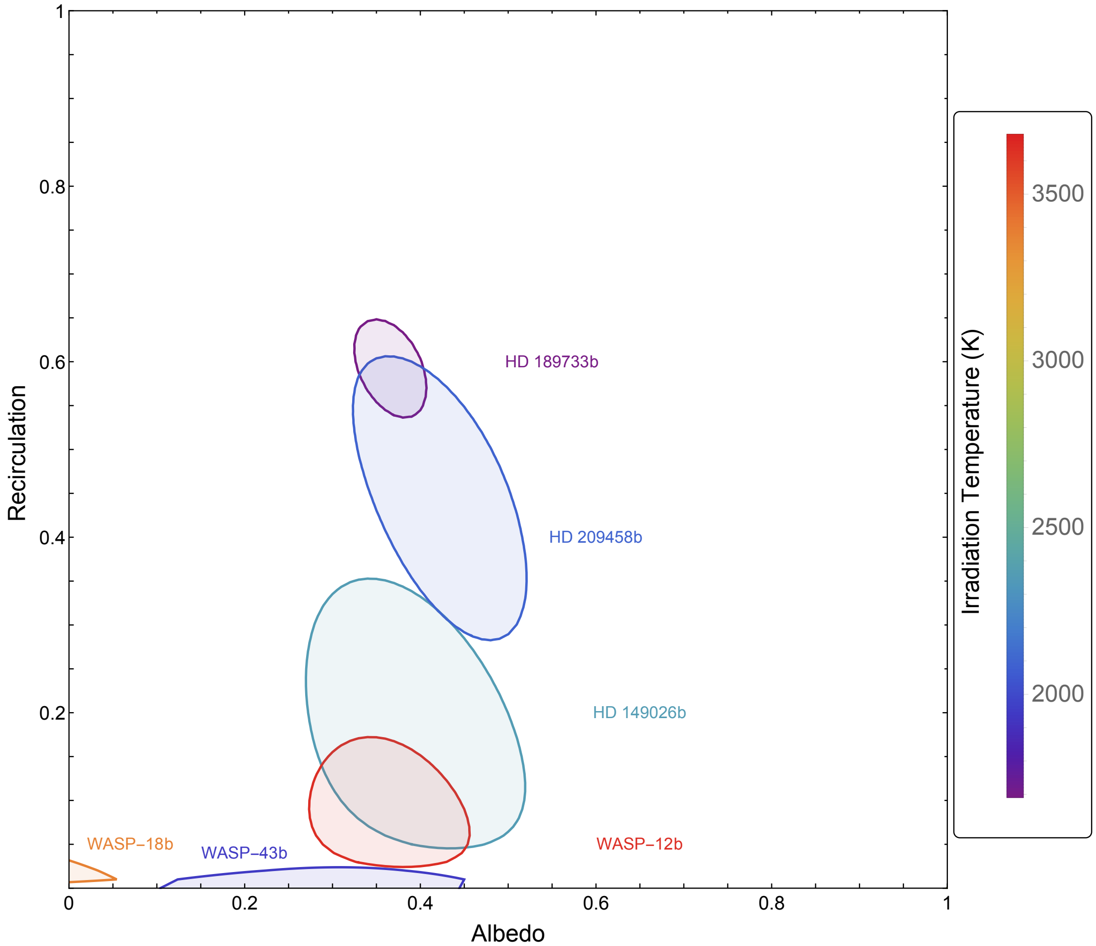


Figure 4.5. Composite 1σ confidence regions for thermal observation planets, as calculated from the error-weighted mean dayside and nightside brightness temperatures. Here the horizontal axis measures *Bond* albedo. Bounding curve colors indicate irradiation temperature: red = warmer, purple = cooler. The inflationary factor $f_{\rm sys} = 3$ is applied to infrared eclipse and phase uncertainties as noted in Table 4.2, but adopting the published eclipse uncertainties barely modifies the confidence intervals.

infer the planet's optical geometric albedo. Our literature review finds nine planets with published eclipse depths at thermal and optical wavelengths, but lacking infrared phase variations (Table 4.2). Note that HD 189733b and HD 209458b benefit from *both* infrared phases and visible eclipses.

Planets in this sample include CoRoT-1b (Alonso, R. et al., 2009a; Zhao et al., 2012; Gillon, M. et al., 2009; Rogers et al., 2013; Deming et al., 2011), CoRoT-2b (Alonso, R. et al., 2009b; Wilkins et al., 2014; Alonso et al., 2010; Deming et al., 2011), HAT-P-7b (Esteves et al., 2014; Christiansen et al., 2010), Kepler-5b (Esteves et al., 2014; Désert et al., 2011b), Kepler-6b (Esteves et al., 2014; Désert et al., 2011b), Kepler-7b (Esteves et al., 2014; Demory et al., 2013), Kepler-13Ab (Esteves et al., 2014; Shporer et al., 2014), TrES-2b (Esteves et al., 2014; Croll et al., 2010a; O'Donovan et al., 2010), and WASP-19b (Abe, L. et al., 2013; Zhou et al., 2013; Bean et al., 2013; Zhou et al., 2014).

4.3.2.1. Thermal Contamination. In order to extract a geometric albedo from an optical eclipse, we must correct the eclipse depth for thermal emission from the planet "leaking" into the visible band (Cowan & Agol, 2011b; Heng & Demory, 2013). In practice, one estimates the planet's dayside effective temperature and extrapolates this into the optical to account for thermal emission at visible wavelengths. However, this procedure is complicated by the fact that real hot Jupiters are vertically and horizontally inhomogeneous, so they emit at higher brightness temperatures in the optical than in the thermal infrared.

Even if every location on a planet emits as a blackbody (BB), the resulting spectrum will not be a Planck curve. For a planet in the zero-albedo and zero-recirculation limit, the equilibrium temperature at any dayside location is described by $T = T_0 \cos^{\frac{1}{4}} \gamma$, where γ is the angle from the sub-stellar point ($\gamma = \frac{\pi}{2}$ at the terminator.) Each annulus of the dayside thus radiates at a different blackbody temperature, and together they produce a "Sum of Blackbodies" (SoB) spectrum (this is analogous to the multicolor blackbody spectra used to model accretion disks; Mitsuda et al., 1984). For fixed bolometric flux, BB

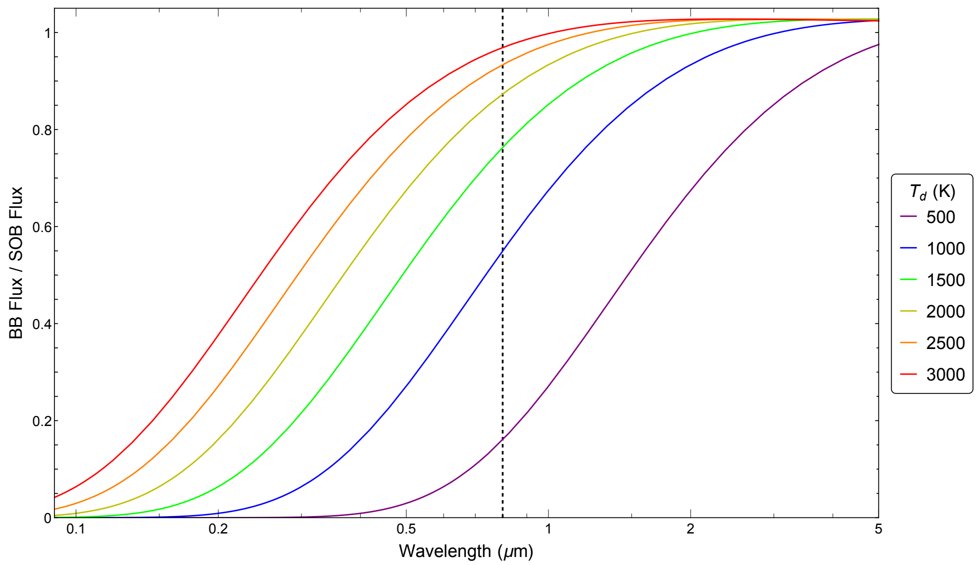


Figure 4.6. Flux ratio between blackbody and “Sum of Blackbodies” spectra for various effective dayside temperatures in 500 K increments. Curves are colored according to temperature: red = warmer, purple = cooler. The vertical dashed line indicates our chosen threshold wavelength, 0.8 μm, between reflected light and thermal emission.

and SoB spectra produce nearly identical flux at thermal wavelengths: the SoB is 1–2 per cent fainter than the BB. At optical wavelengths, however, a BB spectrum underestimates the flux by a factor of a few, as seen in Figure 4.6.

Moreover, the optical photosphere should be deeper and hotter than the infrared photosphere, in a cloud-free atmosphere (Allard et al., 2001; Fortney et al., 2008; Cowan & Agol, 2011a). The combination of these two effects is that a naïve blackbody extrapolation from the infrared to the optical may underestimate thermal emission by a factor of 3–10. In other words, while the hottest planets have the greatest thermal emission at optical wavelengths, somewhat cooler planets have optical emission that is harder to estimate.

Once an optical eclipse has been corrected for thermal contamination, the geometric albedo can be calculated using

$$A_g = \delta_{\rm ecl}^{\rm ref} \left(\frac{a}{R_p} \right)^2, \tag{4.4}$$

where $\delta_{\rm ecl}^{\rm ref}$ is the *reflected light* eclipse depth. Note that geometric albedo is a function of wavelength, while hot Jupiter eclipse depths have typically only been measured in a single optical broadband.

4.3.2.2. Confidence Regions. Our dayside temperature analysis for planets in the eclipse-only sample is analogous to Section 4.3.1.2, and we again perform Monte Carlo simulation with 5000 steps for all uncertainty propagation. For optical eclipses, we optimize computation by first calculating the *uncertainty* of each thermally-corrected geometric albedo, using Equation 4.4 and propagating uncertainties in EWM dayside effective temperature, stellar effective temperature, transit depth, eclipse depth, planetary radius, and semi-major axis. Our thermal correction uses equal contributions of BB and SoB spectra, assuming 10 per cent increase in dayside temperature to account for the vertical temperature profile effect noted in Section 4.3.2.1. To acknowledge variability with this effect, we add a 5 per cent systematic uncertainty in quadrature to the calculated geometric albedo uncertainty. We then construct χ^2 surfaces for each planet as described in Section 4.3.1.2. For optical measurements, we recompute our thermal contamination correction at each χ^2 grid point to determine specific *values* of geometric albedo (since the underlying dayside temperature varies in the albedo–recirculation plane.) Note that HD 189733b has two distinct optical eclipses; we use the weighted mean of both corrected geometric albedos in our χ^2 calculations for this planet.

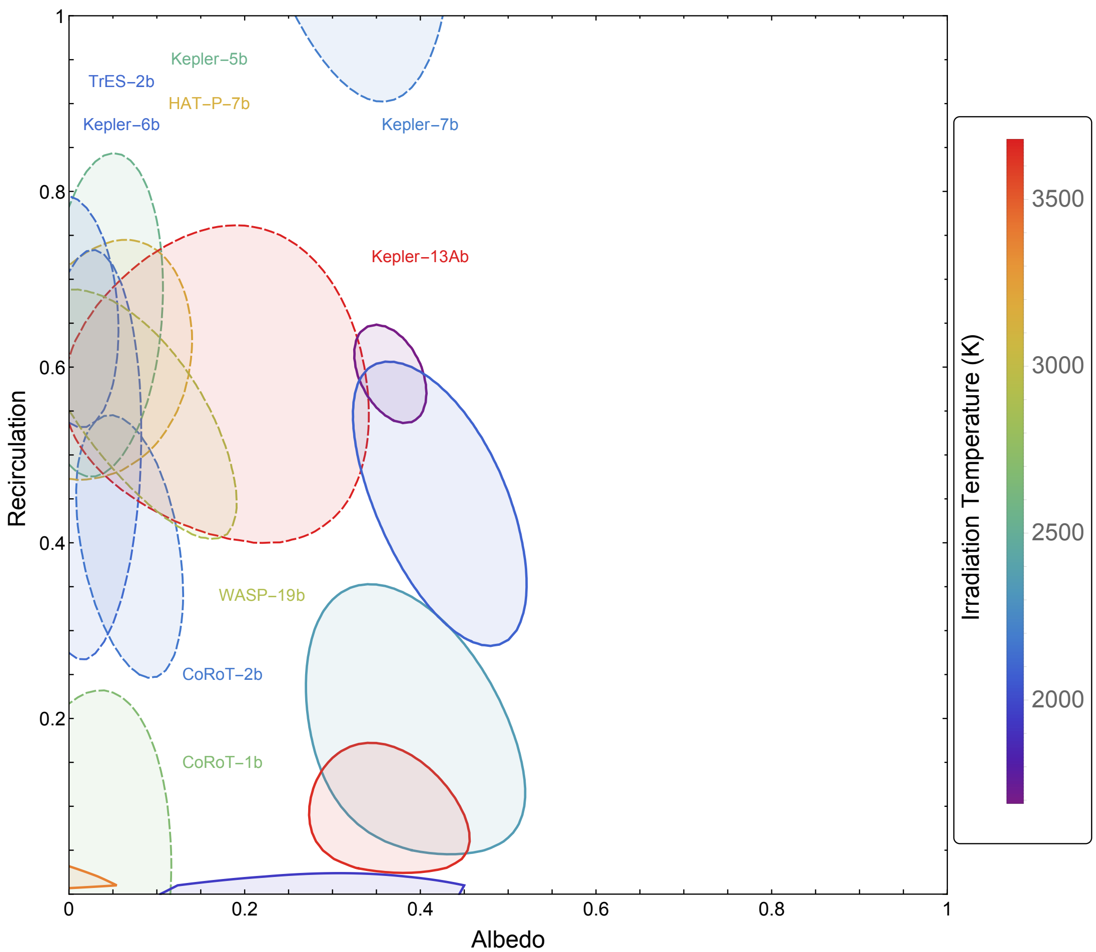


Figure 4.7. Composite 1σ confidence regions for eclipse-only planets as calculated from brightness temperatures (using EWM) and geometric albedos, shown with thermal observation planets for comparison. Here the horizontal axis measures different quantities: *geometric* albedo at visible wavelengths for eclipse-only planets (dashed lines), *Bond* albedo for thermal observation planets (solid lines). Bounding curve color follows Figure 4.5. The inflationary factor $f_{\rm sys} = 3$ is applied to infrared eclipse and phase uncertainties as noted in Table 4.2. Adopting published uncertainties across the board results in similar confidence intervals.

In Figure 4.7 we compare the 1σ confidence intervals of the nine eclipse-only planets to those of the six planets with full-orbit thermal observations. Note that for optical eclipses we implicitly assume $A_B = A_g$, but the actual conversion between geometric and

Table 4.3. Resulting parameters for all planets as calculated from brightness temperatures (via EWM) and geometric albedos, assuming applicable $f_{\rm sys} =$ 3 uncertainty inflation. Thermal observation planets are listed first (with *Bond* albedos), followed by eclipse-only planets (with *geometric* albedos at visible wavelengths.) Low and high values are obtained from the confidence regions of Figures 4.5 and 4.7; fit values are taken to be the grid location of $\chi^2_{\rm min}$.

Planet	Albedo			Recirculation			$\chi^2_{\rm min}$
	Low	**Fit**	High	Low	**Fit**	High	
HD 149026b	0.27	**0.41**	0.519	0.045	**0.15**	0.353	0.0005
HD 189733b	0.325	**0.37**	0.407	0.536	**0.59**	0.648	0.0051
HD 209458b	0.323	**0.43**	0.521	0.283	**0.44**	0.606	0.0013
WASP-12b	0.273	**0.37**	0.456	0.024	**0.07**	0.172	0.0028
WASP-18b	0	**0**	0.054	0.007	**0.01**	0.032	2.2571
WASP-43b	0.104	**0.29**	0.45	0	**0**	0.024	0.7581
CoRoT-1b	0	**0.01**	0.117	0	**0**	0.232	0.0002
CoRoT-2b	0.008	**0.07**	0.13	0.246	**0.4**	0.545	0.0005
HAT-P-7b	0	**0.04**	0.14	0.472	**0.61**	0.745	0.0012
Kepler-5b	0	**0.04**	0.107	0.476	**0.67**	0.844	0.001
Kepler-6b	0	**0.02**	0.082	0.267	**0.51**	0.733	0.005
Kepler-7b	0.258	**0.34**	0.426	0.902	**1**	1	0.9611
Kepler-13Ab	0	**0.18**	0.341	0.4	**0.58**	0.761	0.0004
TrES-2b	0	**0.01**	0.056	0.532	**0.67**	0.795	0.0097
WASP-19b	0	**0.08**	0.191	0.404	**0.56**	0.688	0.0009

Bond albedo is more complicated (Section 4.4.5). Regions are again colored by irradiation temperature, while sample group is denoted by the line style of bounding curve. Though the $f_{\rm sys}$ uncertainty inflation is included for isolated thermal eclipses, taking the published uncertainties at face value produces nearly identical confidence intervals.

Based solely on dayside effective temperatures (Figure 4.2), one might conclude that all planets have roughly the same Bond albedo and heat transport efficiency. Figure 4.7 dispels this notion at high significance. The 1σ intervals for all the planets in Figure 4.7 are listed in Table 4.3: thermal observation planets first, eclipse-only planets second. All best-fit parameters are defined as the location of $\chi^2_{\rm min}$ on the computed grids.

4.4. Discussion

4.4.1. Sources of Error and Uncertainty

It is worth summarizing the various sources of uncertainty and error that we account for in order to produce Figure 4.7. For thermal measurements, we first compute brightness temperatures, accounting for the uncertainties on eclipse depth (inflated by a factor of 3 if based on a single occultation; Hansen et al., 2014), transit depth, and stellar effective temperature, and also compute each planet's irradiation temperature, accounting for uncertainty in stellar effective temperature and scaled semi-major axis. At this stage, we also account for reflected light contamination (non-zero near-infrared geometric albedo). Nightside brightness temperatures are derived the same way, but additionally depend on the thermal phase amplitude and its uncertainty. In converting brightness temperatures to effective temperatures, we account for unknown meridional heat transport and incomplete spectral coverage. The conversion from dayside brightness temperatures to dayside effective temperature is reasonable for hot Jupiters because of their relatively isothermal vertical structure; the conversion may be more fraught for the nightsides of hot Jupiters.

For optical eclipses, we first correct eclipse depths for thermal contamination, accounting for uncertainty on dayside temperature, transit depth, and stellar effective temperature. In this process we also account for vertical and horizontal temperature profiles that conspire to increase optical thermal emission. We next convert the reflected light eclipse depth to an optical geometric albedo, accounting for uncertainties in eclipse depth, planetary radius, and scaled semi-major axis. We assume $A_B = A_g$ for the purposes of constraining heat transport in Figure 4.7 (see Section 4.4.5 for caveats), but this assumption in no way affects our inferred geometric albedo for these planets.

Crucially, for every "correction" that we apply, we add appropriate uncertainty in our inference, either by randomly varying parameters in the Monte Carlo, or by adding systematic uncertainty in quadrature to the formal errors. Our inferences of heat transport, Bond albedo, and geometric albedo are therefore conservative.

4.4.2. Dayside Temperatures

The upward trend in dayside effective temperature with irradiation temperature in Figure 4.2 is unsurprising: one expects highly-irradiated planets to be hotter. The black lines in the plot can be thought of as limiting cases of either heat recirculation or Bond albedo. In the zero-albedo limit, the black lines correspond to $\varepsilon = 0$ (solid), $\varepsilon = 0.4$ (dashed), and $\varepsilon = 1$ (dotted). Alternatively, in the zero-recirculation limit, the black lines correspond to $A_B = 0$ (solid), $A_B = 0.25$ (dashed), and $A_B = 0.625$ (dotted). Therefore, planets that lie above the solid black line must have an internal energy source, while planets lying below the dotted black line must have non-zero Bond albedo.

We can also consider the qualitative claim from Cowan & Agol (2011b) that T_d increases *disproportionately* with T_0. We ignore planets with significantly eccentric orbits, as this complicates their energy budget (these planets are denoted in red in Figure 4.2.) With double the planets and more data per planet, we find that $T_d = -90(80)+0.87(5)T_0$. The χ^2 per datum of the fitted trend is 1.4 ± 0.4, which is a reasonable fit. This is consistent with—but does not strengthen—the claim that planets receiving more stellar flux generally have lower Bond albedo and/or less efficient heat transport.

4.4.3. Thermal Phase Measurements

Figure 4.5 shows a tendency towards lower recirculation efficiency as irradiation increases, in agreement with previous findings (Cowan & Agol, 2011b; Cowan et al., 2012c; Perez-Becker & Showman, 2013). The irradiation temperatures of these planets span approximately 2000 K, corresponding to $\varepsilon = 0.59$ for HD 189733b at the cool end ($T_0 = 1695\ K$) and $\varepsilon = 0.01$ for WASP-18b ($T_0 = 3387\ K$). The irradiation of WASP-12b is actually $\sim 260\ K$ higher than WASP-18b, but their recirculation probability distribution functions overlap (Table 4.3).

The notable exception to the T_0–ε trend is WASP-43b, with $T_0 = 1943\ K$ but exhibiting virtually no heat transport ($\varepsilon = 0$ with $\chi^2_{\rm min} = 0.758$). Our recirculation value is in agreement with the redistribution factor of Stevenson et al. (2014c), and our best-fit Bond albedo ($A_B = 0.29$) is also consistent at the 1σ level. These parameters translate into a cold nightside temperature (nominally $T_n \lesssim 465\ K$), which we routinely find to be consistent with zero. Coupled hydrodynamic and radiative transfer simulations of this planet were able to reproduce its dayside—but not nightside—emission (Kataria et al., 2015), so the poor heat transport of this planet remains a mystery.

The thermal measurements for WASP-18b suggest $A_B \lesssim 0.05$ at 1σ. The planet's best-fit parameters would lie outside the plot to the left, which is indicative of either an internal heat source (identical to a negative Bond albedo in our parametrization) or underestimated observational uncertainties. Kepler-7b and WASP-43b also have $\chi^2_{\rm min}$ well above zero, suggesting that either our model assumptions or the published uncertainties are incorrect. Each of these planets would benefit from more thermal eclipse and phase measurements in order to reduce the systematic uncertainty in dayside and nightside effective temperatures.

HD 189733b and HD 209458b benefit from full-orbit thermal observations as well as optical eclipse measurements, allowing us to compare infrared-based Bond albedos to their optical geometric albedos. For HD 189733b we derive corrected geometric albedos of 0.37(12) at 0.37 μm and 0.04(8) at 0.51 μm, in agreement with Evans et al. (2013). This red-optical geometric albedo is much lower than our Bond albedo estimate of $[0.33, 0.41]$. For HD 209458b we obtain a corrected geometric albedo of 0.04(6) at 0.5 μm, which agrees with Rowe et al. (2008). However, our 1σ interval for Bond albedo is $[0.32, 0.52]$. Therefore, the tentative conclusion based on these two planets is that their Bond albedos are considerably higher than their optical geometric albedos.

4.4.4. Eclipse-Only Measurements

Most of the eclipse-only planets have low geometric albedos: $A_g \lesssim 0.2$ (Figure 4.7). As anticipated, Kepler-7b lies completely above this range (Demory et al., 2011), while the confidence region for Kepler-13Ab extends to a geometric albedo of 0.34. The nine planets exhibit a wide variety of recirculation efficiencies, from CoRoT-1b ($\varepsilon \approx 0.1$) to Kepler-7b ($\varepsilon \approx 0.95$). Eclipse-only planets with similar irradiation temperatures are found at different locations on the ε-axis, and we do not see evidence for a trend between irradiation temperature and recirculation efficiency as with the thermal observation planets. This is not surprising, since dayside measurements offer minimal leverage for inferring the nightside temperature. Strictly speaking we only include these planets in Figure 4.7 by assuming that $A_B = A_g$; the actual comparison is more complex (Section 4.4.5).

Geometric albedo analyses encompassing several planets from our sample have been previously conducted. We compare our results in Table 4.4 to overlapping planets from Heng & Demory (2013): HAT-P-7b, Kepler-5b, Kepler-6b, Kepler-7b, and TrES-2b. Our "uncorrected" geometric albedos for all five planets show agreement within the stated confidence intervals. Esteves et al. (2014) and Angerhausen et al. (2015) also consider these planets, in addition to Kepler-13Ab, under both zero and perfect heat redistribution. Our "full correction" geometric albedos for five of the six planets are in agreement with values from both studies obtained in the maximum equilibrium temperature hypothesis (i.e. hotter dayside temperatures implying greater thermal contamination of the optical eclipse). However, we find Kepler-13Ab to have dissimilar geometric albedo when using stellar parameters from Shporer et al. (2014) with our greater thermal correction: 0.175(113) versus 0.404(55) and $\simeq 0$, respectively. Note this includes stellar effective temperature readjustment of Kepler-13A to $7650 \pm 250\ K$, down from $8500 \pm 400\ K$ (Szabó et al., 2011).

4.4.5. Comparing Geometric to Bond Albedo

In principle, an optical eclipse depth is related to a planet's Bond albedo (Rowe et al., 2006). Indeed, for a range of Solar System planets and moons, the Bond albedo is roughly equal to the optical geometric albedo, albeit with a scatter of ± 30 per cent. Given the possibility of inhomogeneous albedo, uncertainty in the scattering phase function, and unknown reflectance spectrum, it would be imprudent to extrapolate this trend to hot Jupiters. Moreover, Marley et al. (1999) demonstrated that simply varying the incident stellar radiation can alter Bond albedo by a factor of four for identical planets.

Table 4.4. Geometric albedo values, optical waveband starlight fraction, and Bond albedo limits (using Equations 4.6–4.8) for planets with visual eclipses. Observations are denoted by central wavelength and bandwidth. Geometric albedo values are considered as follows: "uncorrected" uses Equation 4.4 assuming no thermal contamination; "simple correction" accounts for thermal emission with a BB spectrum at our error-weighted mean effective dayside temperature; "full correction" uses the method described in Section 4.3.2.2 (equal BB + SoB contributions with vertical temperature profile effect), again using our EWM dayside temperature. Uncertainties on starlight fraction are negligible and thus omitted. All Bond albedo calculations use the "full correction" values and assume the phase integral factor $q = \frac{5}{4}$.

Planet	Wavelength (μm)	Geometric Albedo			$f_*^{\rm opt}$	Bond Albedo		
		Uncorrected	Simple Correction	Full Correction		Min	Gray	High
CoRoT-1b	0.60(42)	0.213(87)	0.129(82)	0.043(77)	0.485	0.026	0.053	0.284
CoRoT-2b	0.60(42)	0.101(35)	0.090(34)	0.069(60)	0.472	0.04	0.086	0.305
HAT-P-7b	0.65(40)	0.262(60)	0.156(54)	0.051(73)	0.438	0.028	0.064	0.309
HD 189733b	0.37(16)	0.374(113)	0.374(113)	0.374(124)	0.091	0.043	0.468	0.497
	0.51(12)	0.042(60)	0.043(61)	0.043(79)	0.123	0.007	0.054	0.445
HD 209458b	0.5(3)	0.039(37)	0.039(36)	0.039(62)	0.384	0.019	0.049	0.327
Kepler-5b	0.65(40)	0.107(22)	0.079(25)	0.040(58)	0.439	0.022	0.05	0.302
Kepler-6b	0.65(40)	0.060(22)	0.047(23)	0.025(55)	0.436	0.014	0.031	0.296
Kepler-7b	0.65(40)	0.314(73)	0.315(72)	0.313(88)	0.44	0.172	0.392	0.452
Kepler-13Ab	0.65(40)	0.468(31)	0.318(67)	0.175(113)	0.405	0.088	0.218	0.386
TrES-2b	0.65(40)	0.030(7)	0.022(8)	0.007(51)	0.439	0.004	0.008	0.284
WASP-19b	0.685(530)	0.248(120)	0.177(113)	0.099(108)	0.537	0.067	0.124	0.298

Our analysis of hot Jupiters suggests that their optical geometric albedos are systematically lower than their Bond albedos. There are three possible explanations for this discrepancy: (1) we have over-corrected the thermal contamination at optical wavelengths, (2) we have systematically underestimated the effective temperatures for planets with full-orbit phase variations, or (3) the geometric albedos of hot Jupiters are, in fact, systematically lower than their Bond albedos because of unexpected scattering phase functions and/or reflectance spectra.

We address the first hypothesis by listing three different geometric albedo calculations in Table 4.4: these differ in their treatment of optical thermal emission. For six of the eleven planets, there is little difference between the geometric albedo estimate after a simple blackbody subtraction as opposed to the scenario with higher optical brightness temperature. For the remaining five planets, the "full correction" geometric albedos are lower than their "simple correction" counterparts. The planets for which the details of thermal emission correction are more important tend to either have higher irradiation temperatures and so greater likelihood for unattributed thermal contamination in the optical (e.g. HAT-P-7b), or have precise optical eclipse measurements where minor changes to the dayside emission have a larger impact on constraining reflected light (e.g. TrES-2b). Our contamination analysis is also largely consistent with the geometric albedos inferred from higher equilibrium temperatures in both Esteves et al. (2014) and Angerhausen et al. (2015). Even in the unlikely event that hot Jupiters have optical dayside brightness temperatures equal to that in the mid-infrared, the optical geometric albedos for most planets are significantly lower than the Bond albedos inferred from thermal phase measurements.

The second solution to the geometric versus Bond albedo discrepancy is that hot Jupiters are much brighter in the NIR, and hence thermal phase measurements—mostly obtained with *Spitzer* in the mid-IR—will underestimate their global temperatures and over-estimate their Bond albedos. However, neither the dayside emission spectra of individual planets, nor their aggregate spectrum, show strong broadband molecular features. This means that dayside brightness temperatures from the near- through mid-IR should be reasonable proxies for their effective temperatures. One may worry that flux is escaping the nightsides of hot Jupiters in the NIR, leading us to underestimate nightside bolometric flux, but that is ruled out for WASP-43b by HST/WFC3 phase measurements, which show *no* nightside flux in the NIR (Stevenson et al., 2014c). It would be useful to have full-orbit NIR phase curves of more planets in order to further test this hypothesis.

This leaves us with the third hypothesis, namely that the Bond albedos of most hot Jupiters are high, despite their low geometric albedos. The geometric albedo of a planet (light reflected back towards the illuminating star) is related to its spherical albedo (light reflected in all directions) by a phase integral, $A_s = qA_g$. Lambertian (diffuse) reflection results in $q = \frac{3}{2}$, while pure Rayleigh scattering produces $q = \frac{4}{3}$. In general, planets with atmospheres—including simulated hot Jupiters—have $1.0 < q < 1.5$ (Pollack et al., 1986; Burrows & Orton, 2010). It would be useful to empirically constrain the scattering phase functions of hot Jupiters using data from space-based photometric missions. In the few cases where reflected phase variations have been measured, the spherical albedo appears so inhomogeneous that it is impossible to infer the phase-dependence of scattering (Demory et al., 2013; Esteves et al., 2014). If we assume that hot Jupiters are diffusely reflecting, then they have typical optical spherical albedos of 15 per cent, still well below the inferred Bond albedos.

The spherical albedo is related to the Bond albedo via a flux-weighted integral (Burrows & Orton, 2010):

$$A_B = \frac{\int_0^\infty A_s(\lambda) I_{\mathrm{inc}}\,\mathrm{d}\lambda}{\int_0^\infty I_{\mathrm{inc}}\,\mathrm{d}\lambda}, \tag{4.5}$$

where I_{inc} is the SED of the incident stellar flux. The degree to which the optical spherical albedo impacts the Bond albedo depends on f_*^{opt}, the percentage of starlight emitted in the observed optical waveband, assuming blackbody radiation at T_* (values of f_*^{opt} are listed in Table 4.4). We consider limiting cases of Equation 4.5, assuming out-of-band wavelengths have average spherical albedos equal to 0, the optical A_s, or 0.5 respectively:

$$A_B^{\mathrm{min}} = A_s f_*^{\mathrm{opt}}, \tag{4.6}$$

$$A_B^{\mathrm{gray}} = A_s, \tag{4.7}$$

$$A_B^{\mathrm{high}} = A_s f_*^{\mathrm{opt}} + 0.5(1 - f_*^{\mathrm{opt}}). \tag{4.8}$$

Our limiting Bond albedos are summarized in Table 4.4. In principle the out-of-band spherical albedo could be unity, but as this would result in Bond albedos so great that the planets would be cooler than is observed in the mid-IR, we adopt a more modest upper threshold in Equation 4.8.

The A_B^{high} scenario is a reasonable match to the Bond albedos inferred from full-orbit thermal measurements. This suggests that most hot Jupiters have geometric albedos of ≈ 50 per cent in the NIR and $\lesssim 10$ per cent in the optical. If hot Jupiters are Lambertian reflectors, the NIR/optical contrast is somewhat less severe. This scenario similar in spirit to the high Bond albedo combined with low red–NIR geometric albedo one can obtain with Rayleigh scattering (Marley et al., 1999), but with the opposite color. Note that the high infrared albedos we are hypothesizing contradict the low infrared geometric albedo we assumed when estimating reflected IR light in Section 4.3.1.2. Using $A_g^{\text{IR}} = 0.5$ in our MC implies greater NIR contamination from reflected starlight, and hence lower dayside thermal flux with greater Bond albedo. The most extreme change is a 20 per cent increase in the Bond albedo of WASP-12b, but nonetheless our conclusions remain unaffected.

It is worth mentioning that geometric albedos of 60 per cent—from the optical through the NIR—were predicted for the hottest giant exoplanets due to reflective silicate clouds (Sudarsky et al., 2000). In order to explain the low optical geometric albedo, one could invoke an optical absorber at low pressures, above the purported cloud deck. Such optical absorbers, originally theorized to explain hot Jupiter temperature inversions, could include gaseous TiO/VO (Fortney et al., 2007) or S_2/HS (Zahnle et al., 2009). In this scenario, Kepler-7b would be unique not because of its clouds, but due to its dearth of optical absorbers.

Alternatively, since cloud reflection is a multiple-scattering process, single-scattering albedos even marginally below unity can result in a heavily muted geometric albedo (Dlugach & Yanovitskij, 1974; Hu, 2014). One might therefore explain the unusually red reflectance spectrum of hot Jupiters with a single cloud deck where individual cloud grains have nearly gray albedo.

Such scenarios might be tested by measuring the optical–infrared transit spectra of hot Jupiters. If the purported absorbers are located at sufficiently low pressures, and if the upper atmospheres of these planets are not too hazy (Pont et al., 2013; Gibson et al., 2013), then we would expect larger effective radii in the optical than the infrared. If instead the red reflectance spectrum is due to multiple scattering within a single cloud deck, then the transit spectrum should be flat.

Moreover, the best way to investigate these hypotheses would be to obtain thermal phase measurements for the planets that have precise optical geometric albedo constraints, and vice versa.

4.5. Updates

4.5.1. Phase Offsets

For transiting planets with *no* atmospheric dynamics and on circular orbits, the observed flux (at some wavelength) looks like the black curve in Figure 4.8. This flux varies during the orbit because one sees different phases of the planet, from the nightside at transit to the dayside at eclipse.

However, phase offsets—the orbital span between center of eclipse and peak planetary brightness—were approximated as zero in previous studies. The dashed orange curve in Figure 4.8 shows an example observation with a phase offset. Here the time of transit and eclipse, and both depths, are identical to the approximation (i.e. black curve), but the peak and trough of the phase curve happen earlier in the orbit. As shown, the difference

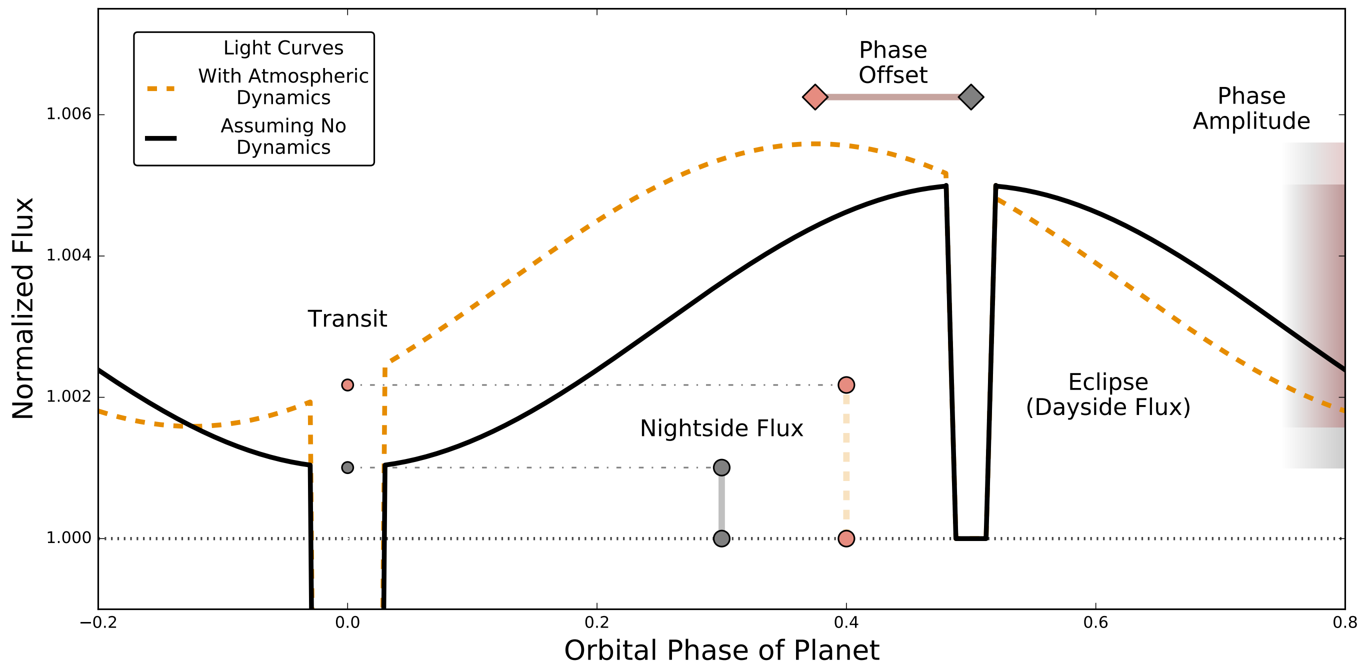


Figure 4.8. Ideal star+transiting planet light curves at an arbitrary wavelength for a given eclipse depth, shown as a function of orbital phase. The dashed orange curve shows a planet with atmospheric dynamics (i.e. reality) that uniformly offsets the *whole* light curve. The black curve instead assumes a planet with no dynamics and so no phase offset—this approximation is used in previous studies (e.g. Schwartz & Cowan, 2015). The transit, eclipse, phase amplitudes, phase offset, and nightside fluxes are all indicated. The stellar flux is normalized to unity and shown as the dotted black line. Both transit depths are the same (bottoms not shown); the ratio of the eclipse and transit depths is an estimate of dayside temperature. For the black curve (e.g. hot Jupiter seen in the infrared), the planet's nightside flux is the difference between the eclipse depth and phase amplitude—dividing by the transit depth gives an estimate of nightside temperature. For a fixed eclipse depth, detecting a phase offset in a light curve means the planet's nightside is hotter relative to its dayside at that wavelength.

in flux between the planet's day and nightside is now smaller than the phase amplitude. Thus, the nightside temperature is higher than before, and so the bulk properties of the planet's atmosphere will be different. We will explore how inferences of Bond albedo and heat recirculation efficiency are affected by phase offsets, using the model of energy balance in Schwartz & Cowan (2015).

4.5.2. Data and Model

We start from the 6 planets in Table 2 of Schwartz & Cowan (2015) that have thermal eclipse and phase data: HD 149026b, HD 189733b, HD 209458b, WASP-12b, WASP-18b, and WASP-43b. Then we update this list with new infrared observations, adding WASP-14b (Wong et al., 2015), HAT-P-7b and WASP-19b (Wong et al., 2016) to our sample.

New eclipses are from Zhou et al. (2015), Evans et al. (2015), and Line et al. (2016, Submitted; WFC3), while phase amplitudes are from Wong et al. (2015), Wong et al. (2016), and Stevenson et al. (2016b). We use the phase offsets in these latter three papers—published offsets for the original sample come from Knutson et al. (2009b,c), Cowan et al. (2012b), Knutson et al. (2012), Maxted et al. (2013), Stevenson et al. (2014c), and Zellem et al. (2014). The 9 planets tested here each have 1–3 measured phase offsets.

In Cowan & Agol (2011b) and Schwartz & Cowan (2015), the intensity of the planet's nightside to its star at some wavelength, $\psi(\lambda)$, is defined as:

$$\psi(\lambda) = \frac{\delta_{\rm ecl} - \delta_{\rm var}}{\delta_{\rm tr}}, \tag{4.9}$$

where $\delta_{\rm ecl}$ is the eclipse depth, $\delta_{\rm var}$ is the phase amplitude, and $\delta_{\rm tr}$ is the transit depth. Because we assume entire light curves are shifted uniformly by phase offsets (i.e. dashed orange curve in Figure 4.8), we modify Equation 4.9 to:

$$\psi'(\lambda) = \frac{\delta_{\rm ecl} - (\delta_{\rm var}\cos\phi_{\rm off})}{\delta_{\rm tr}}, \tag{4.10}$$

where $\phi_{\rm off}$ is the phase offset. For a given eclipse depth, this accounts for a brighter nightside (i.e. Equation 6 of Cowan & Agol, 2011b) when there is an offset, and reduces to Equation 4.9 otherwise.

For each planet we calculate two fits on A_B–ε, using Equation 4.9 or 4.10 in our model. We compile and label these 1σ regions in Figure 4.9. Mimicking Figure 4.8, the light solid curves neglect phase offsets and the dashed curves include them. Each region is colored by irradiation temperature, T_0: this is the hottest the sub-stellar point could be, when the planet absorbs all energy that strikes it and recirculates nothing (Cowan & Agol, 2011b). Note that WASP-43b's solid curve is only on the horizontal axis, just below its dashed region.

In the albedo–recirculation plane, dayside temperature increases to the lower left and nightside temperature towards the upper left. We find, as expected, that the fits with phase offsets (i.e. dashed regions) move in the latter direction. Most nightside effective temperatures increase by $\sim$ 5–20%, and the best-fit Bond albedos drop by up to 0.07. Most recirculation efficiencies rise by up to 0.06—even 0.15 for HD 149026b. The largest relative jump in T_n is for WASP-43b, which climbs $\sim$ 320% to 367 K. WASP-12b has the most significant changes: T_n rises by $\sim$ 50% to 1764 K, A_B drops by 0.12, and ε goes

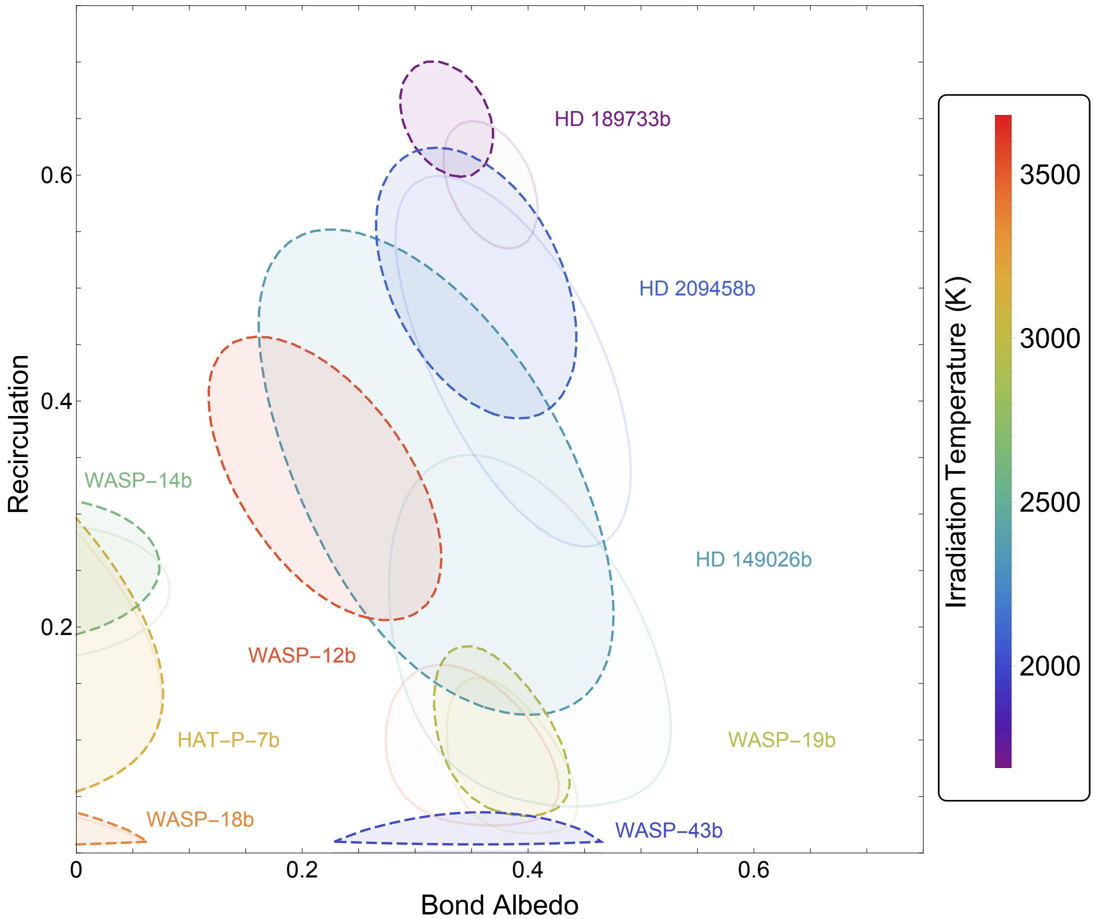


Figure 4.9. The fitted Bond albedo (A_B) and heat recirculation efficiency (ε) for the 9 planets with infrared eclipse and phase data (using χ^2, as in Schwartz & Cowan, 2015; Wong et al., 2015, 2016). Similar to Figure 4.8, the light solid curves show the 1σ regions *without* accounting for phase offsets, while the dashed curves *include* them when available. The solid curve for WASP-43b is a line on the horizontal axis (i.e. nightside has zero flux), below the planet's dashed region. Color shows irradiation temperature—the maximum possible at the sub-stellar point—where red is warmer and purple is cooler. We show up to 0.75 for both parameters, but each can go as high as unity. Adding phase offsets moves a region towards hotter nightside temperature, with lower A_B and higher ε. Note that the different fits for WASP-12b and WASP-43b do not overlap at 1σ. Using phase offsets is preferred because neglecting them, as in previous studies, adds systematic error to the inferences. For a fixed eclipse depth, seeing larger phase offsets in a planet's light curve(s) means that planet absorbs more starlight and moves more heat through its atmosphere from day to night.

up by 0.25. While most of our fits are consistent at 1σ, both WASP-12b and WASP-43b differ significantly depending on how we treat phase offsets.

4.5.3. Previous Results

It is useful to review papers based on the energy balance model in Cowan & Agol (2011b). This study found that (the 24) hot Jupiters globally had $A_B < 0.35$ at 1σ, and those with $T_0 > 2400$ K all had low recirculation efficiency. Perez-Becker & Showman (2013) made a shallow water model to demonstrate the latter: spin synchronized planets have zonal (i.e. East-West) winds and little day-night contrast when irradiation is low, but day-to-night flow with larger differences between T_d and T_n when irradiation is high. Schwartz & Cowan (2015) supported this trend except for WASP-43b (Stevenson et al., 2014c; Kataria et al., 2015), which showed a cold nightside and lower T_0. The thermal and optical data these authors used also showed that (the 50) hot Jupiters had either low or moderate Bond albedos (~ 0.1 vs ~ 0.35)—interpreted as evidence for clouds that reflect infrared light and absorb in the visible.

Wong et al. (2015) next added thermal measurements for WASP-14b (solid green curve in Figure 4.9). WASP-18b had a similarly low Bond albedo, and both were more massive than the other planets modeled (~ 7.7 and ~ 10.1 M_J, vs ~ 1 M_J). Thus, it was suggested these two hot Jupiters could be radiating heat from formation or have stronger Ohmic dissipation. But, Wong et al. (2016) contradicted this with phase data of HAT-P-7b and WASP-19b, since the former has a mass of ~ 1.7 M_J yet very low Bond albedo. Instead the authors stated this could happen if the three planets with low A_B have different thermal evolution histories. We point out that WASP-18b, WASP-14b, and HAT-P-7b have radii within about 1.2–1.5 R_J, roughly 4–14$\times$ closer together (given uncertainties) than the six planets with higher Bond albedos in Figure 4.9.

4.5.4. Impacts

Phase offsets do not significantly change how planets are *distributed* in the A_B–ε plane. We get almost the same probability functions when we marginalize the solid or dashed regions in Figure 4.9 over either axis (not shown), so the studies from Section 4.5.3 are reliable. Yet the fits for *individual* planets are affected, even though the solid and dashed regions usually overlap at 1σ. When we use phase offsets from light curves, our model prefers planets to be less reflective and transport heat better precisely because they have hotter nightside temperatures (Figure 4.8). One should always include phase offsets when studying planetary atmospheres; neglecting them is a first-order approximation.

In fact, the higher nightside temperature we find for WASP-12b ($T_n = 1764 \pm 205$ K) means this planet breaks the model in Perez-Becker & Showman (2013). It receives high stellar irradiation yet has moderate recirculation efficiency; this is opposite to the case of WASP-43b (Stevenson et al., 2014c; Schwartz & Cowan, 2015). Instead, WASP-12b has ε comparable to planets with ~ 1000 K lower T_0 (WASP-14b and HD 149026b). A possible reason is that WASP-12b has a larger radius, $\sim 1.8\ R_J$, and lower density, $\sim 0.24\ \rho_J$ (Hebb et al., 2009), than all other hot Jupiters in our sample, and is likely accreting material onto its host star (e.g. Li et al., 2010). An accretion disk or stream could help transport irradiated gas to cooler parts of the planet's atmosphere—detectable as phase offsets in the light curves. Cowan et al. (2012b) found a larger offset at 3.6 than 4.5 μm, implying hotter material is redistributed more on WASP-12b.

Though it will be extremely difficult to take spectra of Earth analogs with the James Webb Space Telescope, this could be done for particularly close or interesting targets (Beichman et al., 2014). Constraining phase offsets will be crucial to assess habitability over an entire planet. Therefore, light curves of warm terrestrial planets should be carefully analyzed for phase offsets—characterizing climates will depend on it.

4.6. Interlude III

The message from these energy budgets is leading-edge science: this represents some of the best knowledge about the bulk atmospherics of hot Jupiters. To take a step back and think about that for a moment is actually pretty enlightening. It is very common to picture "state of the art" ideas as being unusual or complex, and given the flourish of technology in the last few centuries, that default is definitely well-earned. But even with the monumental effort of building space observatories, finding small changes in stellar brightnesses, and modeling how that light varies over time, the distilled product for our purposes is so inherently modest. There is something uplifting about that—we can help push the entire understanding of planetary atmospheres that float hundreds of light years away with essentially a handful of numbers. As our ability to image terrestrial planets improves, with the James Webb Space Telescope and beyond, this level of detail *will* get finer. The prospect of Exocartography is out over the horizon, where high-precision light curves allow one to deduce surface maps, spin properties, and cloud behavior on distant worlds. Thus, let us take an optimistic step into the future, to illuminate the shroud that will make these features visible someday.

CHAPTER 5

Inferring Planetary Obliquity Using Rotational And Orbital Photometry

This chapter is adapted from Schwartz, J. C., Sekowski, C., Haggard, H. M., Pallé, E., and Cowan, N. B. 2016, MNRAS, 457, 926.

5.1. Introduction

The obliquity of a terrestrial planet encodes information about different processes. A planet's axial alignment and spin rate inform its formation scenario. Numerical simulations have shown that the spin rates of Earth and Mars are likely caused by a few planetesimal impacts (Dones & Tremaine, 1993), while perfect accretion produces an obliquity distribution that is isotropic (e.g. Kokubo & Ida, 2007; Miguel & Brunini, 2010). Conversely, Schlichting & Sari (2007) describe how prograde rotation is preferred to retrograde for a formation model with semi-collisional accretion.

Obliquity is also important in controlling planetary climate. This has been studied in-depth for Earth under many conditions (e.g. Laskar et al., 2004; Pierrehumbert, 2010), and high axial tilts can make planets at large semi-major axes more habitable (Williams & Kasting, 1997). Furthermore, while the Earth's spin axis is stabilized by the Moon (Laskar et al., 1993), obliquities of several Solar System bodies evolve chaotically (Laskar, 1994). This influences searches for hospitable planets, as Spiegel et al. (2009) note that the habitability of terrestrial worlds may depend sensitively on how stable the climate is in the short-term.

A planet's average insolation is set by stellar luminosity and semi-major axis; insolation at different latitudes is determined by obliquity and (for eccentric orbits) the axial orientation. Non-oblique planets have a warmer equator and colder poles that do not vary much throughout the year. Modest obliquities produce seasons at mid-latitudes because the sub-stellar point moves North and South during the orbit (Pierrehumbert, 2010). Planets tilted at angles $\geq 54°$ receive more overall radiation near their poles and have large orbital variations in temperature (Williams & Pollard, 2003). Thus, even limited knowledge of a planet's obliquity can help constrain the spatial dependence of insolation and temperature.

Numerous methods have been proposed for measuring planetary obliquities. Seager & Hui (2002) and Barnes & Fortney (2003) demonstrated constraints on oblateness and obliquity using ingress/egress differences in transit light curves; Carter & Winn (2010) extended and applied these techniques to observations of HD 189733b. Kawahara (2012) derived constraints on obliquity from modulation of a planet's radial velocity during orbit, while Nikolov & Sainsbury-Martinez (2015) examined the Rossiter-McLauglin effect at secondary eclipse for transiting exoplanets. One could also measure obliquity at infrared wavelengths, using polarized rotational light curves (De Kok et al., 2011) and orbital variations (e.g. Gaidos & Williams, 2004; Cowan et al., 2013).

A planet's obliquity can also be constrained by changes in reflected light, which will be studied with forthcoming optical and near-infrared space missions, such as *ATLAST* (Postman et al., 2010), *LUVOIR* (Kouveliotou et al., 2014), and *HDST* (Dalcanton et al., 2015). Time-resolved measurements of a rotating planet in one photometric band can reveal its rotation rate (Ford et al., 2001; Pallé et al., 2008; Oakley & Cash, 2009); this helps determine Coriolis forces and predict large-scale circulation. Multi-band photometry can reveal colors of clouds and surface features (Ford et al., 2001; Fujii et al., 2010, 2011; Cowan & Strait, 2013), and enables a longitudinal albedo map to be inferred from disk-integrated light (Cowan et al., 2009). High-cadence, reflected light measurements spanning a full planetary orbit constrain a planet's obliquity and two-dimensional albedo map (Kawahara & Fujii, 2010, 2011; Fujii & Kawahara, 2012).

However, these results have not yet been established for lower cadence measurements of non-terrestrial planets. We also hope to establish a conceptual understanding of how photometric measurements constrain obliquity. While this is less immediately practical, a deeper understanding of this inversion has the potential to lead to further advances in inferring planetary geometry from limited data sets. In this paper, we study light curve methods for arbitrary albedo maps and viewing geometries, and demonstrate they are useful even for observations at only one or two orbital phases.

Light curves of planets encode the viewing geometry and hence a planet's obliquity because different latitudes are impinged by starlight at different orbital phases (this "kernel" is described in Section 5.2.2). To see this, consider a planet with no obliquity in an edge-on, circular orbit. The star always illuminates the Northern and Southern hemispheres equally, and we never view some latitudes more than others. If instead this planet were tilted, the Northern hemisphere would be lit first, then the Southern hemisphere half an orbit later. If the planet is not North-South uniform, its apparent albedo (Qui et al., 2003; Cowan et al., 2009) would change during its orbit, shown in the left panel of Figure 5.1.

One may also learn about a planet's obliquity as it rotates. Imagine a zero-obliquity planet in a face-on, circular orbit: the observer always sees the Northern pole with half the longitudes illuminated. For an oblique planet, however, more longitudes would be lit when the visible pole leans towards the star, and vice versa. Zero-obliquity planets in edge-on orbits are similar, since more longitudes are lit near superior conjunction, or fullest phase. If the planet has East-West albedo variations, then this longitudinal width will modulate the apparent albedo of the planet as it spins, shown in the right panel of Figure 5.1.

Our work is organized as follows: in Section 5.2, we summarize the observer viewing geometry and explain the reflective kernel, both in two- and one-dimensional forms. Section 5.3.1 introduces a case study planet and describes the kernel at single orbital phases; we consider time evolution in Section 5.3.2. We discuss real observations in Section 5.4.1, then develop our case study in Sections 5.4.2–5.4.4, demonstrating that even single- and dual-epoch observations could allow one to constrain obliquity. In Section 5.4.5, we discuss how to distinguish a planet's rotational direction by monitoring its apparent albedo. Section 5.5 summarizes our conclusions. For interested readers, a full mathematical description of the illumination and viewing geometry is presented in Appendix B.1. Details about the kernel and its relation to a planet's apparent albedo are described in Appendix B.2.

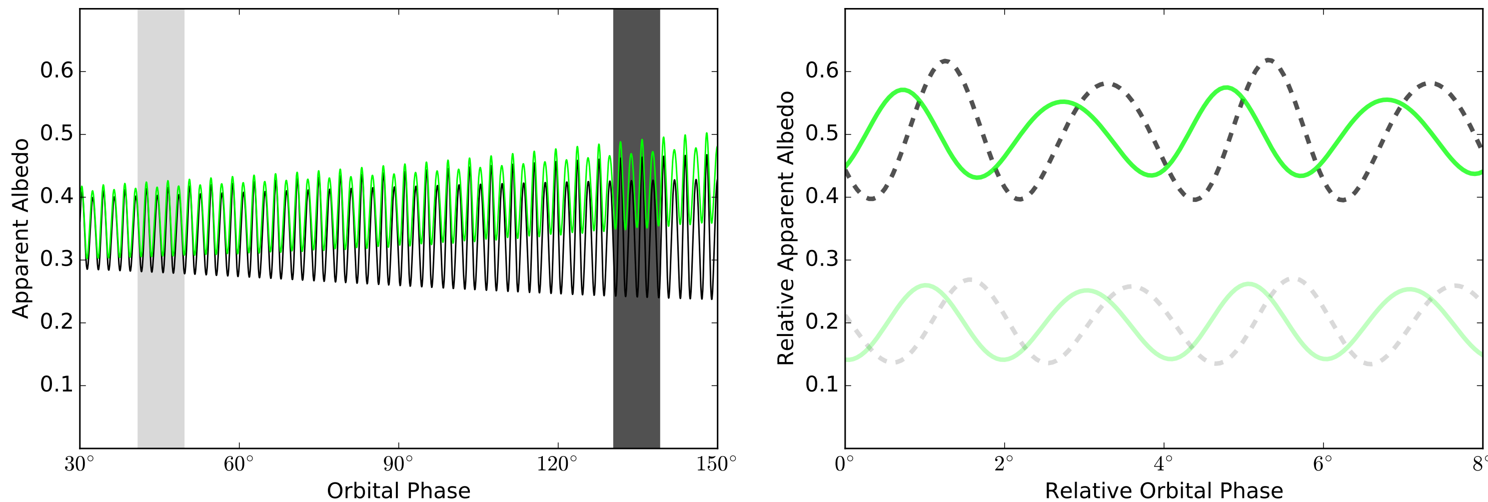


Figure 5.1. The left panel shows apparent albedo as a function of orbital phase for an arbitrary planet with North-South and East-West albedo markings, seen edge-on with zero obliquity (black) or 45° obliquity (green). The average albedo of the green planet increases during the orbit because brighter latitudes become visible and illuminated; this does not happen for the black planet. The vertical bands are each roughly two planetary days, enlarged at right, where lighter shades are the fuller phase. For clarity, the rotational curves are shifted and the zero-obliquity planet is denoted by a dashed line. The apparent albedo of both planets varies more over a day when a narrower range of longitudes and albedo markings are visible and illuminated, and vice versa. Note that both light curves are produced by the *same* albedo map, and are distinct solely because of differences in the kernel for these two geometries. Orbital and/or rotational changes in apparent albedo can help one infer a planet's obliquity.

5.2. Reflected Light

5.2.1. Geometry & Flux

The locations on a planet that contribute to the disk-integrated reflected light depend only on the sub-observer and sub-stellar positions, which both vary in time. A complete development of this viewing geometry is provided in Appendix B.1, which we summarize here. We neglect axial precession and consider planets on circular orbits. Assuming a static albedo map, the reflected light seen by an observer is determined by the colatitude and longitude of the sub-stellar and sub-observer points, explicitly θ_s, ϕ_s, θ_o, and ϕ_o. The intrinsic parameters of the system are the orbital and rotational angular frequencies, $\omega_{\rm orb}$ and $\omega_{\rm rot}$ (where positive $\omega_{\rm rot}$ is prograde), and the planetary obliquity, $\Theta \in [0, \pi/2]$. Extrinsic parameters differ from one observer to the next; we denote the orbital inclination, i (where $i = 90^\circ$ is edge-on), and solstice phase, $\xi_{\rm s}$ (the orbital phase of Summer solstice for the Northern hemisphere). We also define initial conditions for orbital phase, ξ_0, and the sub-observer longitude, $\phi_o(0)$. Reflected light is then completely specified by these seven parameters and the planet's albedo map.

We consider only diffuse (Lambertian) reflection in our analysis. Specular reflection, or glint, can be useful for detecting oceans (Williams & Gaidos, 2008; Robinson et al., 2010, 2014), but is a localized feature and a minor fraction of the reflected light at gibbous phases. The reflected flux measured by a distant observer is therefore a convolution of the two-dimensional kernel (or weight function; Fujii & Kawahara, 2012), $K(\theta, \phi, \mathbb{S})$, and the planet's albedo map, $A(\theta, \phi)$:

$$F(t) = \oint K(\theta, \phi, \mathbb{S}) A(\theta, \phi) \mathrm{d}\Omega, \tag{5.1}$$

where F is the observed flux, θ and ϕ are colatitude and longitude, and $\mathbb{S} \equiv \{\theta_s, \phi_s, \theta_o, \phi_o\}$ implicitly contains the time-dependencies in the sub-stellar and sub-observer locations. Reconstructing a map of an exoplanet based on time-resolved photometry can be thought of as a deconvolution (Cowan et al., 2013), while estimating a planet's obliquity amounts to backing out the kernel of the convolution.

The sub-stellar and sub-observer points are completely determined through a function $\mathbb{S} = f(\mathbb{G}, \omega_{\rm rot} t)$, where $\mathbb{G} \equiv \{\xi(t), i, \Theta, \xi_{\rm s}\}$ is the system geometry and $\xi(t)$ is orbital phase. This is made explicit in Appendix B.1. For a planet with albedo markings, one would therefore fit the observed flux to infer both the planet's albedo map (Cowan et al., 2009) and spin axis (Kawahara & Fujii, 2010, 2011; Fujii & Kawahara, 2012). To study how these methods work for arbitrary maps and geometries, we will focus on the kernel from Equation 5.1, which we can analyze independent of a planet's albedo map.

5.2.2. Kernel

The kernel combines illumination and visibility, defined for diffuse reflection in Cowan et al. (2013) as

$$K(\theta, \phi, \mathbb{S}) = \frac{1}{\pi} V(\theta, \phi, \theta_o, \phi_o) I(\theta, \phi, \theta_s, \phi_s), \tag{5.2}$$

where $V(\theta, \phi, \theta_o, \phi_o)$ is the visibility and $I(\theta, \phi, \theta_s, \phi_s)$ is the illumination. Visibility and illumination are each non-zero over one hemisphere at any time, and are further given by

$$\begin{aligned} V(\theta, \phi, \theta_o, \phi_o) = \max\big[&\sin\theta \sin\theta_o \cos(\phi - \phi_o) \\ &+ \cos\theta\cos\theta_o, 0\big], \end{aligned} \tag{5.3}$$

$$I(\theta, \phi, \theta_s, \phi_s) = \max\Big[\sin\theta \sin\theta_s \cos(\phi - \phi_s) + \cos\theta \cos\theta_s, 0\Big]. \tag{5.4}$$

As noted above, we can express the kernel equivalently as

$$K(\theta, \phi, \mathbb{S}) = K(\theta, \phi, \mathbb{G}, \omega_{\mathrm{rot}} t), \tag{5.5}$$

though we will drop the rotational dependence for now because it does not affect our analysis. We return to rotational frequency in Section 5.4.5.

The non-zero portion of the kernel is a lune: the illuminated region of the planet that is visible to a given observer. The size of this lune depends on orbital phase, or the angle between the sub-observer and sub-stellar points. A sample kernel is shown at the top of Figure 5.2, where the purple and yellow contours are visibility and illumination, respectively. The peak of the kernel is marked with an orange diamond.

We begin by calculating time-dependent sines and cosines of the sub-observer and sub-stellar angles for a viewing geometry of interest (Appendix B.1). These are substituted into Equations 5.3 and 5.4 to determine visibility and illumination at any orbital phase. The two-dimensional kernel is then calculated on a 101 $\times$ 201 grid in colatitude and longitude.

5.2.3. Longitudinal Width

The two-dimensional kernel, $K(\theta, \phi, \mathbb{G})$, is a function of latitude and longitude that varies with time and viewing geometry. For observations with minimal orbital coverage or planets that are uniform from North to South, different latitudes are hard to distinguish

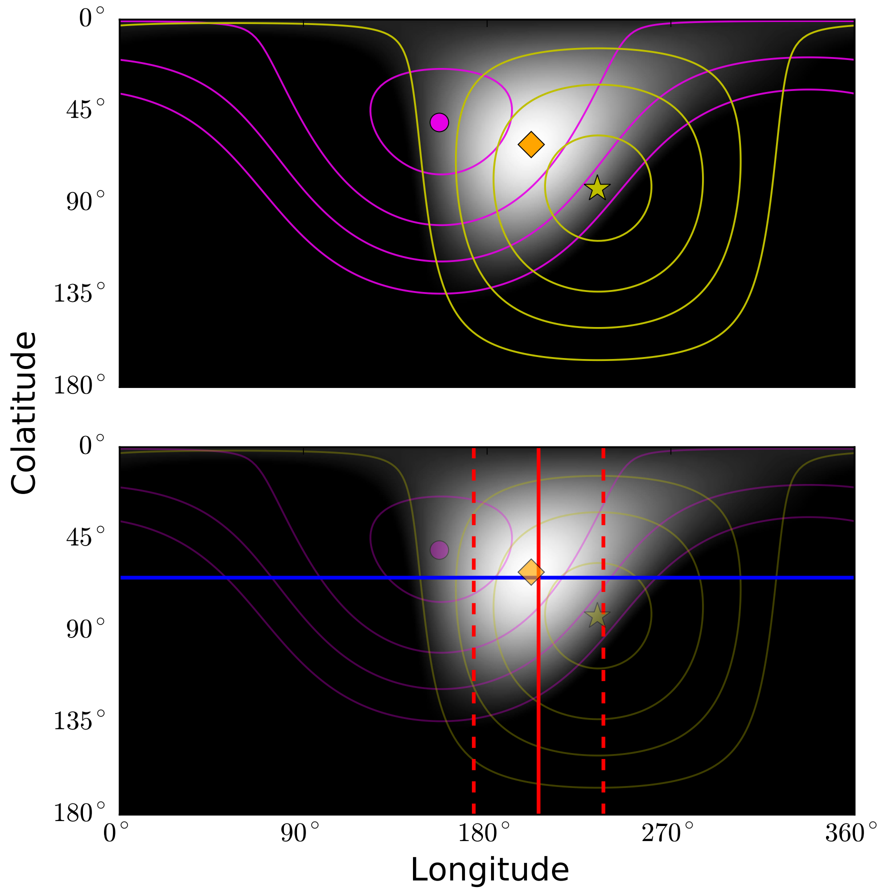


Figure 5.2. *Upper:* A kernel, in gray, with contours showing visibility and illumination, in purple and yellow, as in Cowan et al. (2013). The sub-observer and sub-stellar points are indicated by the purple circle and yellow star, respectively. The orange diamond marks the peak of the kernel. *Lower:* The mean of the longitudinal kernel and the width from this mean are shown as solid and dashed red lines; the dominant colatitude is shown as a blue line.

apart (Cowan et al., 2013) and we use the longitudinal form of the kernel, $K(\phi, \mathbb{G})$, given by

$$K(\phi, \mathbb{G}) = \int_0^\pi K(\theta, \phi, \mathbb{G}) \sin\theta \mathrm{d}\theta. \tag{5.6}$$

We can approximately describe $K(\phi, \mathbb{G})$ by a longitudinal mean, $\bar{\phi}$, and width, σ_ϕ. These are defined in Appendix B.2.1; examples are shown as vertical red lines in the bottom panel of Figure 5.2.

For any geometry, we can calculate the two-dimensional kernel and the corresponding longitudinal width. The mean longitude is unimportant by itself because, for now, we are only concerned with the size of the kernel. We compute a four-dimensional grid of kernel widths with 5° resolution in orbital phase (time), inclination, obliquity, and solstice phase. The result is $\sigma_\phi(\xi(t), i, \Theta, \xi_\mathrm{s}) \equiv \sigma_\phi(\mathbb{G})$, and our numerical grid has size $73 \times 19 \times 19 \times 73$ in the respective parameters. Example contours from this array at first quarter phase, or $\xi(t) = 90°$, are shown in the left panels of Figure 5.3. In these plots obliquity is radial: the center is $\Theta = 0°$ and the edge is $\Theta = 90°$. The azimuthal angle gives the orientation (solstice phase) of the planet's obliquity.

5.2.4. Dominant Colatitude

For a given planet and observer, the sub-observer colatitude is fixed but the sub-stellar point moves North and South throughout the orbit if the planet has non-zero obliquity. This means different orbital phases will probe different latitudes, as dictated by the kernel. To analyze these variations we use the latitudinal form of the kernel, $K(\theta, \mathbb{G})$, explicitly:

$$K(\theta, \mathbb{G}) = \int_0^{2\pi} K(\theta, \phi, \mathbb{G}) \mathrm{d}\phi. \tag{5.7}$$

We may describe $K(\theta, \mathbb{G})$ by its dominant colatitude (Cowan et al., 2012a), $\bar{\theta}$, also defined in Appendix B.2.1 and shown as a horizontal blue line at the bottom of Figure 5.2. We produce a four-dimensional dominant colatitude array, $\bar{\theta}(\xi(t), i, \Theta, \xi_\mathrm{s}) \equiv \bar{\theta}(\mathbb{G})$, similarly

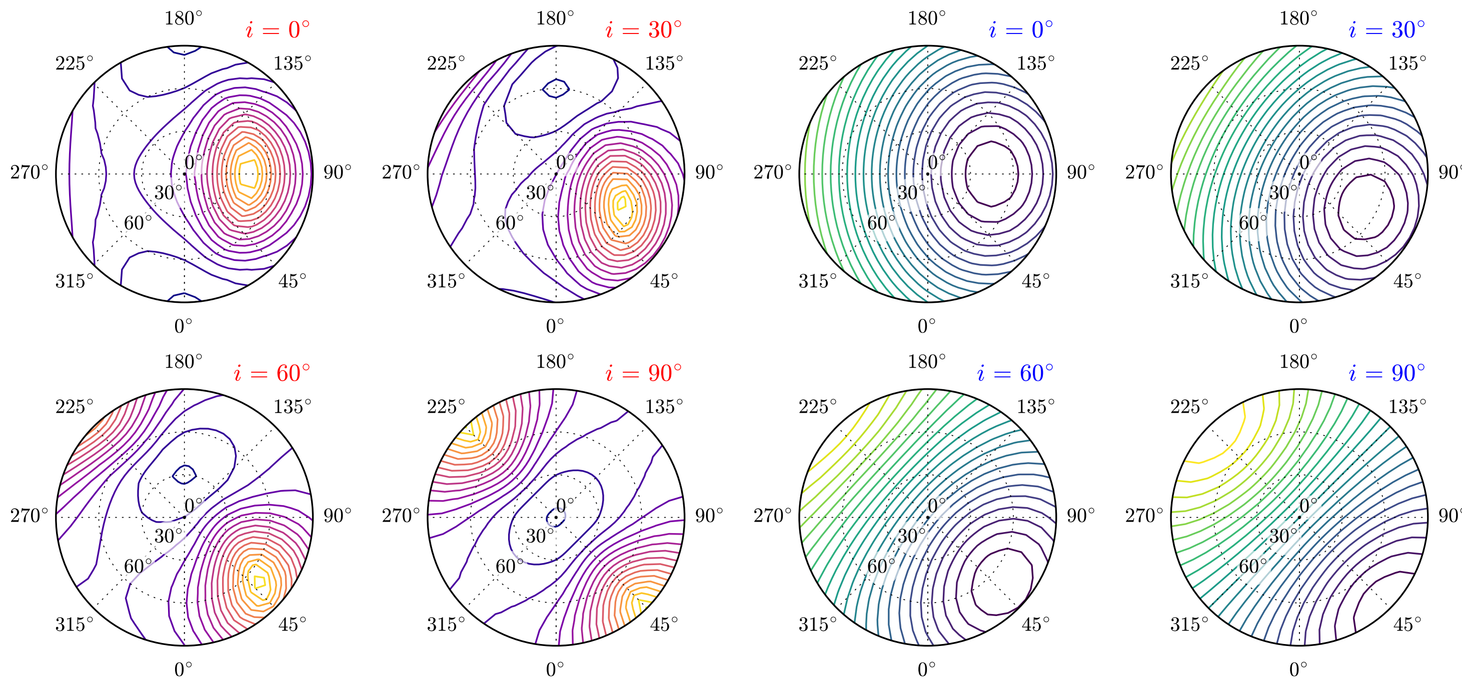


Figure 5.3. *Left panels:* Contours of longitudinal kernel width at first quarter phase, $\xi(t) = 90°$, as a function of planetary obliquity, from face-on $i = 0°$ to edge-on $i = 90°$. Obliquity is plotted radially: the center is $\Theta = 0°$ and the edge is $\Theta = 90°$. The azimuthal angle represents the orientation (solstice phase) of the planet's obliquity. The contours span 20°–100°, dark to light colors, in 5° increments. Larger kernel widths—and hence muted rotational variability—occur at gibbous phases, and when the peak of the kernel is near a pole. *Right panels:* Analogous contours of dominant colatitude that span 35°–145°. A lower dominant colatitude—and thus more reflection from the Northern hemisphere—occurs when the kernel samples Northern regions more, and vice versa.

to $\sigma_\phi(\mathbb{G})$ from Section 5.2.3. Sample contours from this array at first quarter phase are shown in the right panels of Figure 5.3.

5.3. Kernel Behavior

We now consider how the longitudinal width and dominant colatitude of the kernel depend on a planet's obliquity. As a case study, we will define the inclination and spin axis of a hypothetical planet, Q:

$$\text{planet } Q \equiv \left\{ \begin{array}{c} i = 60^\circ \\ \\ \Theta = 55^\circ \\ \\ \xi_\mathrm{s} = 260^\circ \end{array} \right\}. \tag{5.8}$$

5.3.1. Phases

Considering a single orbital phase defines a three-dimensional slice through $\sigma_\phi(\mathbb{G})$ and $\bar{\theta}(\mathbb{G})$ that describes the kernel at that specific time. We show the longitudinal form of the kernel for planet Q at different phases in the left panel of Figure 5.4. Lighter colors are fuller phases, indicating the kernel narrows as this planet orbits towards inferior conjunction, or $\xi(t) = 180^\circ$. The kernel width influences the rotational light curve at a given phase: narrower kernels can have larger amplitude variability in apparent albedo on a shorter timescale (e.g. right of Figure 5.1).

The latitudinal kernel for planet Q is shown similarly in the right panel of Figure 5.4. We see that the kernel preferentially probes low and mid-latitudes during the first half-orbit. The dominant colatitude of planet Q, indicated by circles, also fluctuates during this portion of the orbit—and eventually shifts well into the Northern hemisphere after

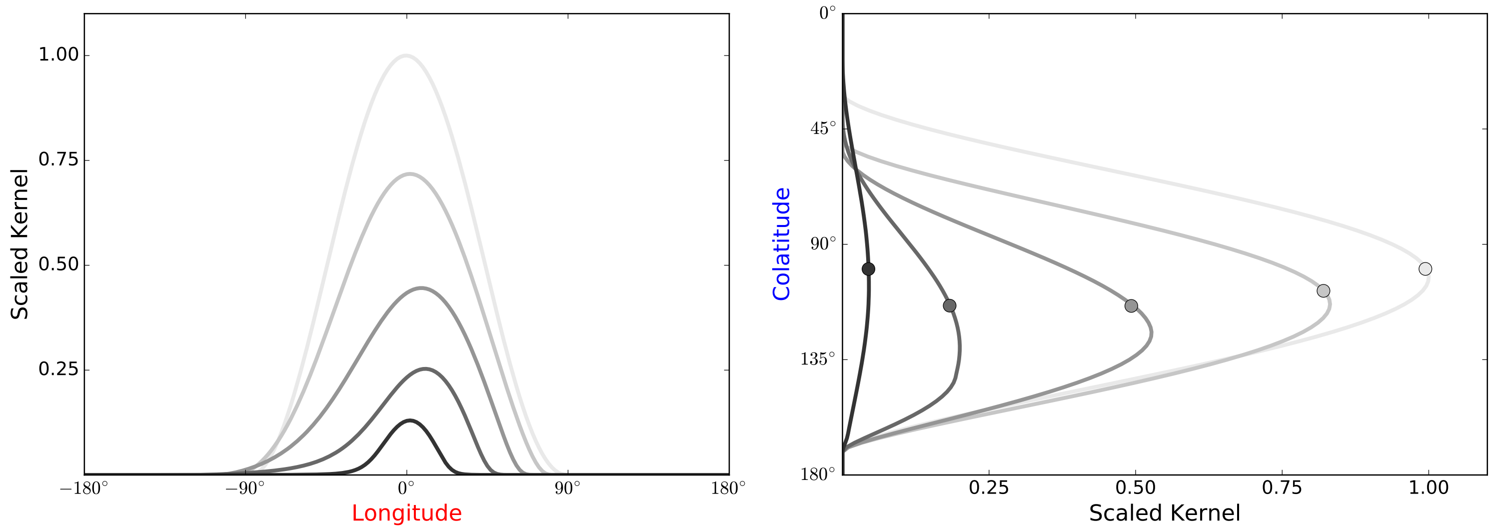


Figure 5.4. *Left:* Longitudinal kernel for planet Q, defined in Equation 5.8, at orbital phases from 30° to 150°, light to dark shades, in 30° increments. Values are scaled to the maximum of the lightest curve. Longitude is measured from each kernel mean. The kernel width decreases as this planet approaches inferior conjunction, or as color darkens. *Right:* Analogous latitudinal kernel for planet Q, where the dominant colatitude, indicated by a circle, increases then returns towards the equator.

inferior conjunction (not shown). Note that the dominant colatitude is *not* always at the peak of the latitudinal kernel (see also Figure B.2). Since one needs measurements at multiple phases to be sensitive to latitudinal variations in albedo, we will consider *changes* in dominant colatitude from one phase to the next, $|\Delta\bar{\theta}|$, for planets with North-South albedo markings. Larger changes in dominant colatitude can make the apparent albedo vary more between orbital phases (e.g. left of Figure 5.1).

5.3.2. Time Evolution

Kernel width and dominant colatitude both vary throughout a planet's orbit. We investigate this by slicing $\sigma_\phi(\mathbb{G})$ and $\bar{\theta}(\mathbb{G})$ along obliquity and/or solstice phase. To start, we vary planet Q's obliquity and track kernel width as shown in the left panel of Figure 5.5. The actual planet Q is denoted by a dashed green line: this planet has a narrow kernel width during the first half-orbit that widens sharply after inferior conjunction. The largest variations between the traces occur near $\xi(t) \approx \{120°, 210°\}$.

We also show tracks of dominant colatitude in the right panel of Figure 5.5. What matters is the change in this characteristic between two epochs; diverse changes in the traces occur between $\xi(t) \approx \{135°, 240°\}$. Planet Q is again the dashed green line, and near the middle of all the tracks more often than for kernel width. If one has some prior knowledge of the viewing geometry, then Figure 5.5 implies at which phases one could observe to best distinguish obliquities for planet Q—for example, $\xi(t) \approx \{120°, 240°\}$.

We can instead vary the solstice phase of planet Q while keeping its obliquity fixed (not shown). In most cases, solstice phase impacts the kernel width and dominant colatitude

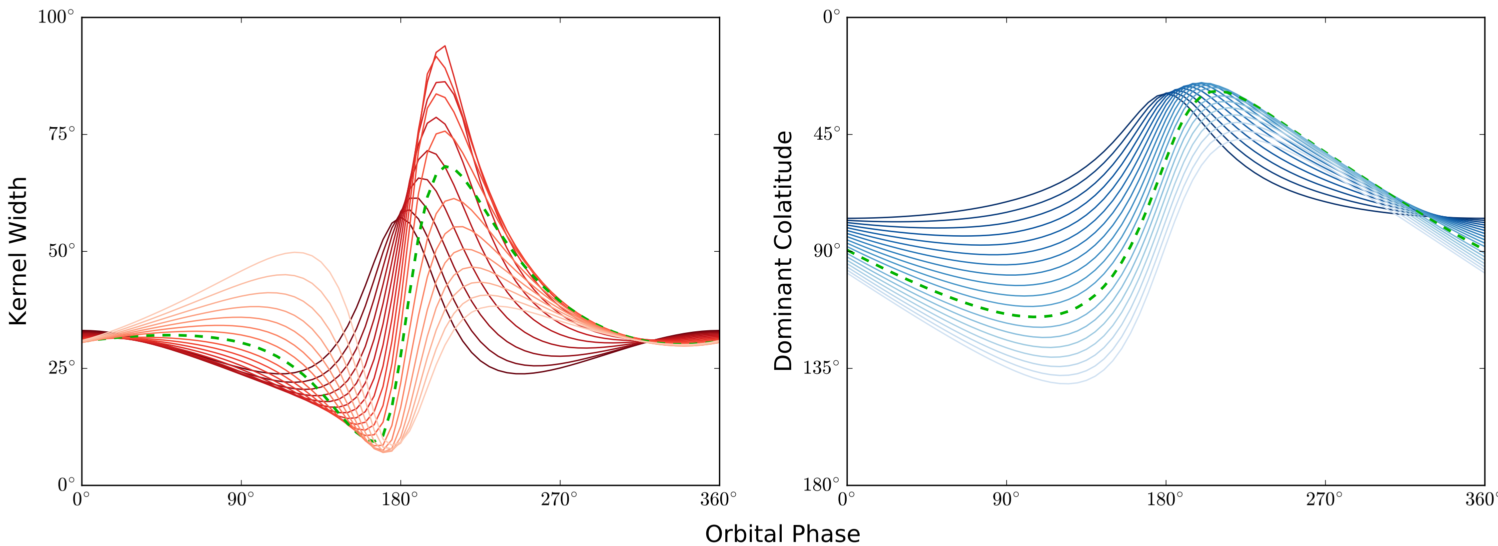


Figure 5.5. *Left:* Kernel width for planet Q as a function of orbital phase, with obliquity varied in 5° increments. The defined planet Q obliquity, $\Theta = 55°$, is the dashed green line; darker and lighter shades of red denote obliquities closer to 0° and 90°, respectively. Inferior conjunction occurs at $\xi(t) = 180°$. The largest variations are near $\xi(t) \approx \{120°, 210°\}$. *Right:* Analogous dominant colatitude for planet Q. Dual-epoch changes are diverse between $\xi(t) \approx \{135°, 240°\}$, for example.

as much as the axial tilt. This is expected, since obliquity is a vector quantity with both magnitude and orientation.

5.4. Discussion

5.4.1. Observations

By analyzing the kernel, we can learn how observed flux may depend on a planet's obliquity independent from its albedo map. For real observations, one would fit the light curve to directly infer the planet's albedo map (Cowan et al., 2009) and spin axis (Kawahara & Fujii, 2010, 2011; Fujii & Kawahara, 2012). We will use the kernel to *predict* how single- and dual-epoch observations constrain planetary obliquity. We address our assumptions below.

The planetary inclination and orbital phase of observation must be known to model the light curve accurately. Both angles might be obtained with a mixture of astrometry on the host star (e.g. *SIM PlanetQuest*; Unwin et al., 2008), direct-imaging astrometry (Bryden, 2015), and/or radial velocity. We will assume that inclination and orbital phase have each been measured with 10° uncertainty.

Extracting the albedo map and spin axis from a light curve could also be difficult in practice. Planets with completely uniform albedo are not amenable to these methods. Moreover, one cannot distinguish latitudes for planets that are North-South uniform, nor longitudes for those that are East-West uniform. Even if a planet has albedo contrast, photometric uncertainty adds noise to the reflected light measurements. Contrast ratios $\leq 10^{-11}$ are needed to resolve rotational light curves of an Earth-like exoplanet (Pallé et al., 2008), which should be achievable by a *TPF*-type mission with high-contrast coronagraph or starshade (Ford et al., 2001; Trauger & Traub, 2007; Turnbull et al., 2012; Cheng-Chao et al., 2015).

We will implicitly assume that planet Q has both East-West and North-South albedo markings, and thus that the kernel geometry impacts the reflected light. In particular, we will assume two scenarios: perfect knowledge of the kernel, or kernel widths and changes in dominant colatitude that are constrained to $\pm 10^{\circ}$ and $\pm 20^{\circ}$, respectively, explained in Appendix B.2.2. Note that these uncertainties will depend on the planet's albedo contrast, and the photometric precision, in a non-linear way. We envision a triage approach for direct-imaging missions: planets that vary in brightness the most, and thus have the easiest albedo maps and spin axes to infer, will be the first for follow-up observations.

Of course, planetary radii are necessary to convert fluxes into apparent albedos (Qui et al., 2003; Cowan et al., 2009). Radii will likely be unknown, but could be approximated using mass-radius relations and mass estimates from astrometry or radial velocity, or inferred from bolometric flux using thermal infrared direct-imaging (e.g. *TPF-I*; Beichman et al., 1999; Lawson et al., 2008). Real planets may also have variable albedo maps, e.g. short-term variations from changing clouds and smaller variations from long-term seasonal changes (Robinson et al., 2010), that could influence the apparent albedo on orbital timescales. These are difficulties that will be mitigated with each iteration of photometric detectors and theoretical models.

5.4.2. Longitudinal Constraints

A fit to the rotational light curve can be used to constrain the spin axis (and longitudinal map) of a planet with East-West albedo contrast. We can demonstrate these constraints using kernel widths in two ways, described in Section 5.4.1 and shown in the upper panels of Figure 5.6. The dark dashed lines and square are idealized constraints when assuming perfect knowledge of the orbital geometry and two kernel widths: $\sigma_{\phi 1} = 25.2°$ at $\xi(t_1) = 120°$, and $\sigma_{\phi 2} = 51.7°$ at $\xi(t_2) = 240°$. Alternatively, the red regions have $10°$ uncertainty on each width (Appendix B.2.2), where we use a normalized Gaussian probability density and include Gaussian weights for uncertainties on inclination and orbital phase.

The green circles represent the true planet Q spin axis, which always lies on the idealized constraints. Only two spin axes are consistent with the ideal kernel widths from both orbital phases. We run more numerical experiments for a variety of system geometries (not shown) and find that perfect knowledge of the kernel width at three orbital phases uniquely determines the planetary spin axis. However, we find a degeneracy for planets with edge-on orbits, where two different spin configurations will produce the same kernel widths at all phases.

As anticipated, we also find planet Q's spin axis (green circle) consistent with the uncertain kernel widths (dark red regions). Imperfect kernel widths at two phases allow all obliquities above $15°$ at 1σ, but exclude nearly one-fifth of spin axes at 3σ. We find similar predictions for other orbital phases and planet parameters.

These examples also suggest that obliquity could be constrained for planets with variable albedo maps. As long as albedo only changes on timescales longer than the rotational period, light curves will constrain both the instantaneous map and the planet's spin axis. A given spin orientation and orbital inclination dictates a specific kernel width as a function of orbital phase (left panel of Figure 5.5), so we predict that light curves at three

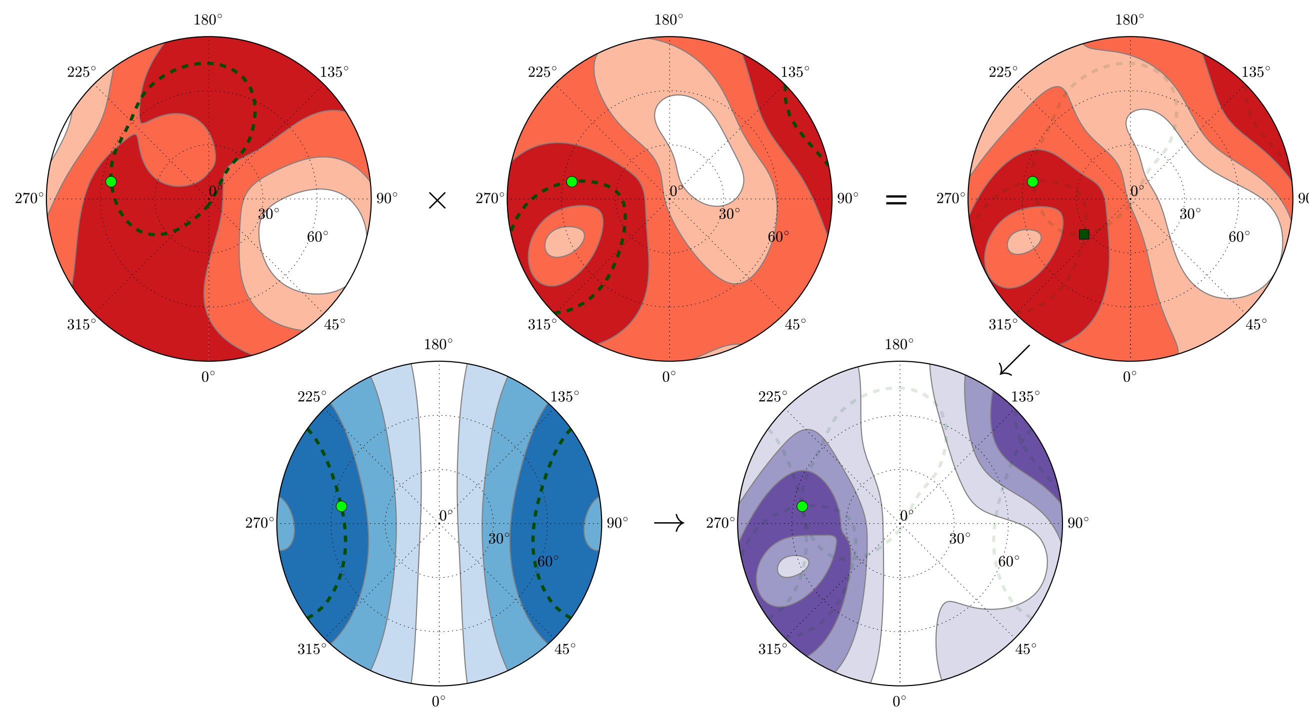


Figure 5.6. Predicted confidence regions for planet Q's spin axis, from hypothetical single- and dual-epoch observations. The constraints are predicted using the kernel, as described in Sections 5.4.1–5.4.4 and Appendix B.2.2: longitudinal widths in the upper row (red), the change in dominant colatitude at the lower left (blue), and joint constraints at the lower right (purple). Obliquity is plotted radially: the center is $\Theta = 0°$ and the edge is $\Theta = 90°$. The azimuthal angle represents the planet's solstice phase. The green circles are the true planet Q spin axis, while the dark dashed lines and square show idealized constraints assuming perfect knowledge of the orbital geometry and kernel (i.e. no uncertainties). The upper left and center panels describe planet Q at $\xi(t) = \{120°, 240°\}$, respectively, while the lower left panel incorporates both phases. For the colored regions, 10° uncertainty is assumed on each kernel width, inclination, and orbital phase, while 20° uncertainty is assumed on the change in dominant colatitude. Regions up to 3σ are shown, where darker bands are more likely. Observing a planet at just a few orbital phases can significantly constrain both its obliquity and axial orientation.

phases will be sufficient to pin down the planetary obliquity, even if the planet's map varies between phases.

5.4.3. Latitudinal Constraints

A fit to light curves from different orbital phases can be used to constrain the spin axis (and latitudinal map) of a planet with North-South albedo contrast. As described in Section 5.4.1, we can demonstrate this constraint using both perfect and uncertain knowledge of the change in dominant colatitude. Our predictions are shown in the lower left panel of Figure 5.6. The idealized constraint here is $|\Delta\bar{\theta}_{12}| = 76.0^\circ$ between $\xi(t) = \{120^\circ, 240^\circ\}$. Since the change in dominant colatitude is constrained between pairs of epochs, four orbital phases are needed to produce three independent constraints and uniquely determine planet Q's spin axis. We find the same two-fold degeneracy as before for planets in edge-on orbits, even if one knows the change in dominant colatitude between all pairs of phases.

For the blue regions, we reapply our probability density from above and assume 20° uncertainty on the change in dominant colatitude (Appendix B.2.2). The distribution is bimodal because only the *magnitude* of the change can be constrained, not its direction. This means an observer would not know whether more Northern or Southern latitudes are probed at the later phase, affecting which spin axes are inferred.

5.4.4. Joint Constraints

For a planet with both East-West and North-South albedo contrast, one may combine longitudinal and latitudinal information to better constrain the planet's spin axis (and two-dimensional map). We show this for planet Q in the lower right panel of Figure 5.6.

The idealized constraint shows that only the true spin configuration is allowed. We find this result for other system geometries—except using orbital phases 180° apart, which creates a two-way degeneracy in the spin axis.

The confidence regions in purple assume our notional uncertainties on both kernel width and change in dominant colatitude (Appendix B.2.2). This prediction is not unimodal, but the 1σ region excludes obliquities below 30°. A distant observer would know that this planet's obliquity has probably not been eroded by tides (Heller et al., 2011), and that the planet likely experiences obliquity seasons.

5.4.5. Pro/Retrograde Rotation

The sign of rotational angular frequency (positive = prograde) can affect the mean longitude of the kernel, but not its size and shape. There is a formal degeneracy for edge-on, zero-obliquity cases: prograde planets with East-oriented maps have identical light curves to retrograde planets with West-oriented maps. The motion of the kernel peak is the same over either version of the planet, implying the retrograde rotation in an inertial frame is *slower* (Appendix B.2.3). We show this scenario in the left panel of Figure 5.7, where the dashed brown line is the difference in prograde and retrograde apparent albedo. The orange and black planets are always equally bright because the same map features, in the upper panels, are seen at the same times.

However, the spin direction of oblique planets and/or those on inclined orbits may be deduced. Inclinations that are not edge-on most strongly alter a planet's light curve near inferior conjunction, seen in the center panel of Figure 5.7: this planet's properties are intermediate between the edge-on, zero-obliquity planet and planet Q. While a typical

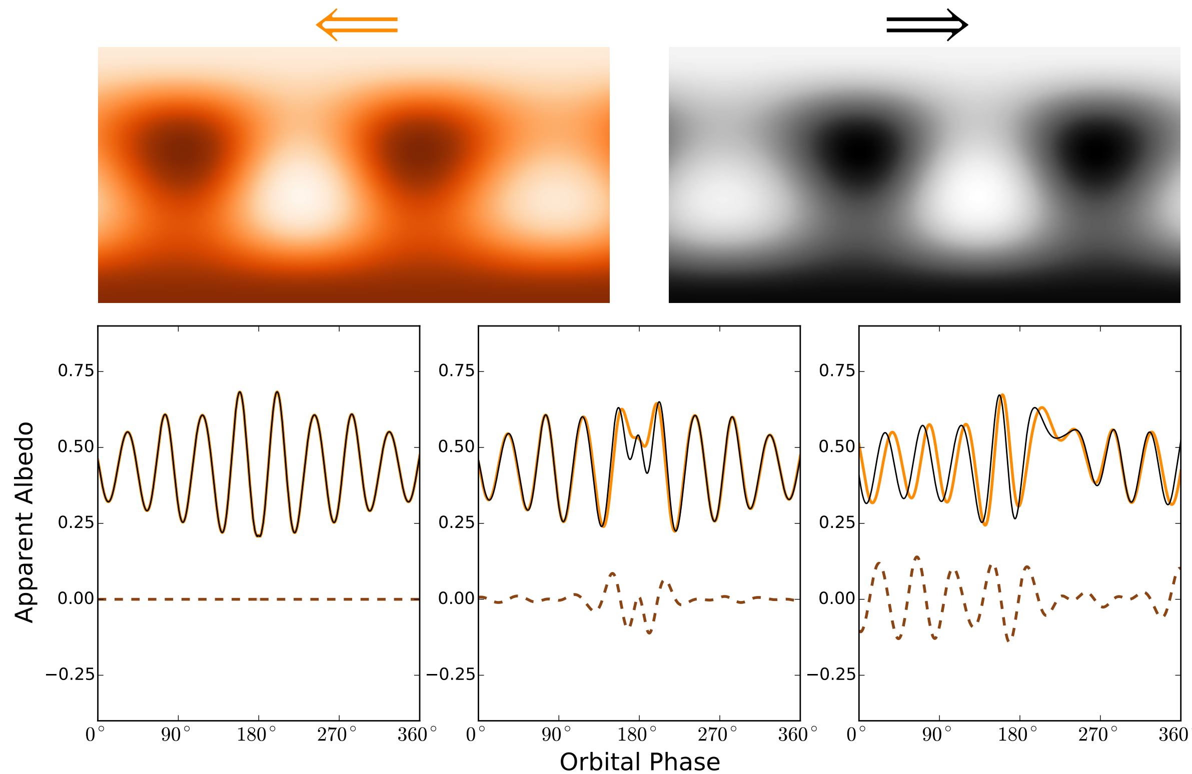


Figure 5.7. Apparent albedo as a function of orbital phase, shown for an edge-on, zero-obliquity planet on the left, planet Q on the right, and an intermediate planet in the center. A low rotational frequency is used for clarity. The black and orange curves correspond to prograde and retrograde rotation, respectively, and the differences in apparent albedo are the dashed brown lines. Inferior conjunction occurs at $\xi(t) = 180°$. The albedo maps are color-coded at the top, where arrows indicate spin direction and the prime meridians are centered. Note that these maps are East-West reflections of each other. The edge-on, zero-obliquity curves are identical, while the curves for the intermediate planet and planet Q grow more distinct. Edge-on, zero-obliquity planets are hopeless, but one can distinguish pro/retrograde rotation for inclined, oblique planets by monitoring their brightness, particularly near crescent phases.

observatory's inner working angle would hide some of the signal, differences on the order of 0.1 in the apparent albedo would be detectable at extreme crescent phases. Alternatively, higher obliquity causes deviations that—depending on solstice phase—can arise around one or both quarter phases. This happens for planet Q in the right panel of Figure 5.7, where both effects combine to distinguish the spin direction at most phases.

Inclination and obliquity influence apparent albedo because the longitudinal motion of the kernel peak is *not* the same at all latitudes. One can break this spin degeneracy in principle, but we have not fully explored the pro/retrograde parameter space. In general, the less inclined and/or oblique a planet is, the more favorable crescent phases are for determining its spin direction.

5.5. Conclusions

We have performed numerical experiments to study the problem of inferring a planet's obliquity from time-resolved photometry, for arbitrary albedo maps and viewing geometries. We have demonstrated that a planet's obliquity will influence its light curve in two distinct ways: one involving East-West albedo markings and another involving North-South markings. Provided this planet is not completely uniform, one could constrain both its albedo map and spin axis using reflected light.

The kernel—the product of visibility and illumination—has a peak, a longitudinal width, and a mean latitude that vary in time and are functions of viewing geometry. Analyzing the kernel enables us to predict constraints on a planet's spin axis from reflected light, including for maps that are East-West uniform (e.g. Jupiter-like) or North-South uniform (e.g. beach ball-like). Curiously, we find that kernel width offers better constraints on obliquity than dominant colatitude, suggesting that East-West albedo contrast is generally more useful than North-South contrast. This is partly because kernel width can be constrained even for variable albedo maps.

Furthermore, monitoring a planet at only a few epochs could determine its spin direction and significantly constrain its obliquity. In our case study of planet Q, we find crescent phases are favorable for telling prograde from retrograde rotation. Similarly, perfect knowledge of the kernel width at two orbital phases narrows the possible spin axes for planet Q to two distinct configurations, while kernel width uncertainties of 10° still exclude about three-quarters of spin axes at 1σ. Adding the constraint on change in dominant colatitude between the same two phases completely specifies the true spin configuration of planet Q. A change in dominant colatitude with 20° uncertainty excludes five-sixths of spin orientations at 1σ.

Most importantly, we also find that perfect knowledge of the kernel width at just three phases, or its change in dominant colatitude between four phases, is generally sufficient to uniquely determine a planet's obliquity. This suggests that—in principle—rotational light curves at 2–4 distinct orbital phases uniquely constrain the spin axis of any planet with non-uniform albedo. This is good news for inferring the obliquity of planets with future direct-imaging missions.

References

Abe, L. et al. 2013, A&A, 553, A49

Agol, E., Cowan, N. B., Knutson, H. A., Deming, D., Steffen, J. H., Henry, G. W., & Charbonneau, D. 2010, ApJ, 721, 1861, 1007.4378

Agol, E., Steffen, J., Clarkson, W., et al. 2005, Monthly Notices of the Royal Astronomical Society, 359, 567

Akaike, H. 1974, Automatic Control, IEEE Transactions on, 19, 716

Allard, F., Hauschildt, P. H., Alexander, D. R., Tamanai, A., & Schweitzer, A. 2001, The Astrophysical Journal, 556, 357

Alonso, R., Deeg, H., Kabath, P., & Rabus, M. 2010, The Astronomical Journal, 139, 1481

Alonso, R. et al. 2009a, A&A, 506, 353

Alonso, R., Guillot, T., Mazeh, T., Aigrain, S., Alapini, A., Barge, P., Hatzes, A., & Pont, F. 2009b, A&A, 501, L23

Anderson, D. et al. 2010, Astronomy & Astrophysics, 513, L3

——. 2011, Monthly Notices of the Royal Astronomical Society, 416, 2108

Anderson, D. P., Cobb, J., Korpela, E., Lebofsky, M., & Werthimer, D. 2002, Communications of the ACM, 45, 56

Angerhausen, D., DeLarme, E., & Morse, J. A. 2015, Publications of the Astronomical Society of the Pacific, 127, 1113

Anglada-Escudé, G. et al. 2016, Nature, 536, 437

Baglin, A., Auvergne, M., Barge, P., Michel, E., Catala, C., Deleuil, M., & Weiss, W. 2007in , AIP Publishing, 201–209

Bailey, J. 2014, PASA - Publications of the Astronomical Society of Australia, 31

Ballard, S. et al. 2010, Publications of the Astronomical Society of the Pacific, 122, 1341

——. 2011, The Astrophysical Journal, 743, 200

Baraffe, I., Selsis, F., Chabrier, G., Barman, T. S., Allard, F., Hauschildt, P., & Lammer, H. 2004, Astronomy & Astrophysics, 419, L13

Barbe, D. F. 1975, Proceedings of the IEEE, 63, 38

Barclay, T. et al. 2013, Nature, 494, 452

Barman, T. S., Hauschildt, P. H., & Allard, F. 2005, The Astrophysical Journal, 632, 1132

Barnes, J., Barman, T. S., Prato, L., Segransan, D., Jones, H., Leigh, C., Cameron, A. C., & Pinfield, D. 2007, Monthly Notices of the Royal Astronomical Society, 382, 473

Barnes, J. W., & Fortney, J. J. 2003, The Astrophysical Journal, 588, 545

Barnes, R. et al. 2016, arXiv preprint arXiv:1608.06919

Barnes, R., Mullins, K., Goldblatt, C., Meadows, V. S., Kasting, J. F., & Heller, R. 2013, Astrobiology, 13, 225

Baskin, N. J. et al. 2013, The Astrophysical Journal, 773, 124

Batalha, N. M. et al. 2013, The Astrophysical Journal Supplement Series, 204, 24

Batygin, K., & Brown, M. E. 2016, The Astronomical Journal, 151, 22

Bean, J. L., Désert, J.-M., Seifahrt, A., Madhusudhan, N., Chilingarian, I., Homeier, D., & Szentgyorgyi, A. 2013, The Astrophysical Journal, 771, 108

Beatty, T. G. et al. 2014, The Astrophysical Journal, 783, 112

Beaulieu, J.-P. et al. 2006, Nature, 439, 437

——. 2011, The Astrophysical Journal, 731, 16

Beerer, I. M. et al. 2010, The Astrophysical Journal, 727, 23

Beichman, C. et al. 2014, Publications of the Astronomical Society of the Pacific, 126, 1134

Beichman, C. A., Woolf, N., & Lindensmith, C. 1999, The Terrestrial Planet Finder (TPF): a NASA Origins Program to search for habitable planets/the TPF Science Working Group; [Washington, DC]: NASA (JPL publication; 99-3)

Biller, B. A. et al. 2013, The Astrophysical Journal Letters, 778, L10

Blecic, J. et al. 2014, The Astrophysical Journal, 781, 116

——. 2013, The Astrophysical Journal, 779, 5

Borovikov, S., & Pogorelov, N. 2014, The Astrophysical Journal Letters, 783, L16

Borucki, W. J. et al. 2013, Science, 340, 587

——. 2011, The Astrophysical Journal, 736, 19

Boss, A. P. 2000, The Astrophysical Journal Letters, 536, L101

Bouchy, F. et al. 2005, Astronomy & Astrophysics, 444, L15

Bowler, B. P. 2016, arXiv preprint arXiv:1605.02731

Broeg, C. et al. 2013, in EPJ Web of Conferences, Vol. 47, EDP Sciences, 03005

Bryden, G. 2015, IAU General Assembly, 22, 58195

Burrows, A., Budaj, J., & Hubeny, I. 2008, The Astrophysical Journal, 678, 1436

Burrows, A., Hubeny, I., Budaj, J., Knutson, H., & Charbonneau, D. 2007, The Astrophysical Journal Letters, 668, L171

Burrows, A., & Orton, G. 2010, Exoplanets, 419

Burrows, A. S. 2014, Proceedings of the National Academy of Sciences, 201304208

Burrows, C. J., Holtzman, J. A., Faber, S., Bely, P. Y., Hasan, H., Lynds, C., & Schroeder, D. 1991, The Astrophysical Journal, 369, L21

Butler, R. P., Marcy, G. W., Williams, E., Hauser, H., & Shirts, P. 1997, The Astrophysical Journal Letters, 474, L115

Campo, C. J. et al. 2011, The Astrophysical Journal, 727, 125

Carter, J. A., & Winn, J. N. 2009, The Astrophysical Journal, 704, 51

——. 2010, The Astrophysical Journal, 709, 1219

Chambers, J. E. 2004, Earth and Planetary Science Letters, 223, 241

Chan, T., Ingemyr, M., Winn, J. N., Holman, M. J., Sanchis-Ojeda, R., Esquerdo, G., & Everett, M. 2011, The Astronomical Journal, 141, 179

Charbonneau, D. et al. 2005, The Astrophysical Journal, 626, 523

——. 2009, Nature, 462, 891

Charbonneau, D., Brown, T. M., Burrows, A., & Laughlin, G. 2006, arXiv preprint astro-ph/0603376

Charbonneau, D., Knutson, H. A., Barman, T., Allen, L. E., Mayor, M., Megeath, S. T., Queloz, D., & Udry, S. 2008, The Astrophysical Journal, 686, 1341

Chauvin, G. et al. 2012, Astronomy & Astrophysics, 542, A41

Chen, G., van Boekel, R., Madhusudhan, N., Wang, H., Nikolov, N., Seemann, U., & Henning, T. 2014, Astronomy & Astrophysics, 564, A6

Chen, J., & Chen, Z. 2008, Biometrika, 95, 759

Cheng-Chao, L., De-Qing, R., Jian-Pei, D., Yong-Tian, Z., Xi, Z., Gang, Z., Zhen, W., & Rui, C. 2015, Research in Astronomy and Astrophysics, 15, 453

Christiansen, J. L. et al. 2010, The Astrophysical Journal, 710, 97

Ciardi, D. R., Fabrycky, D. C., Ford, E. B., Gautier III, T., Howell, S. B., Lissauer, J. J., Ragozzine, D., & Rowe, J. F. 2013, The Astrophysical Journal, 763, 41

Cochran, W. D., Hatzes, A. P., & Hancock, T. J. 1991, The Astrophysical Journal, 380, L35

Cowan, N., Agol, E., & Charbonneau, D. 2007, Monthly Notices of the Royal Astronomical Society, 379, 641

Cowan, N., Deming, D., Gillon, M., Knutson, H., Madhusudhan, N., & Rauscher, E. 2011, Spitzer Proposal, 1, 80044

Cowan, N. et al. 2015, arXiv preprint arXiv:1502.00004

Cowan, N. B., Abbot, D. S., & Voigt, A. 2012a, The Astrophysical Journal Letters, 752, L3

Cowan, N. B., & Agol, E. 2008, ApJL, 678, L129, 0803.3622

Cowan, N. B., & Agol, E. 2011a, The Astrophysical Journal, 726, 82

——. 2011b, The Astrophysical Journal, 729, 54

Cowan, N. B. et al. 2009, The Astrophysical Journal, 700, 915

Cowan, N. B., Fuentes, P. A., & Haggard, H. M. 2013, Monthly Notices of the Royal Astronomical Society, 434, 2465

Cowan, N. B., Machalek, P., Croll, B., Shekhtman, L. M., Burrows, A., Deming, D., Greene, T., & Hora, J. L. 2012b, The Astrophysical Journal, 747, 82

Cowan, N. B., & Strait, T. E. 2013, The Astrophysical Journal Letters, 765, L17

Cowan, N. B., Voigt, A., & Abbot, D. S. 2012c, The Astrophysical Journal, 757, 80

Croll, B. et al. 2014, arXiv preprint arXiv:1410.4286

Croll, B., Albert, L., Lafreniere, D., Jayawardhana, R., & Fortney, J. J. 2010a, The Astrophysical Journal, 717, 1084

Croll, B., Jayawardhana, R., Fortney, J. J., Lafrenière, D., & Albert, L. 2010b, The Astrophysical Journal, 718, 920

Croll, B., Lafreniere, D., Albert, L., Jayawardhana, R., Fortney, J. J., & Murray, N. 2010c, The Astronomical Journal, 141, 30

Crossfield, I. J., Barman, T., Hansen, B. M., Tanaka, I., & Kodama, T. 2012a, The Astrophysical Journal, 760, 140

Crossfield, I. J., Hansen, B. M., Harrington, J., Cho, J. Y.-K., Deming, D., Menou, K., & Seager, S. 2010, The Astrophysical Journal, 723, 1436

Crossfield, I. J., Knutson, H., Fortney, J., Showman, A. P., Cowan, N. B., & Deming, D. 2012b, The Astrophysical Journal, 752, 81

Crouzet, N., McCullough, P. R., Burke, C., & Long, D. 2012, The Astrophysical Journal, 761, 7

Crouzet, N., McCullough, P. R., Deming, D., & Madhusudhan, N. 2014, The Astrophysical Journal, 795, 166

Cubillos, P., Harrington, J., Madhusudhan, N., Foster, A. S., Lust, N. B., Hardy, R. A., & Bowman, M. O. 2014, The Astrophysical Journal, 797, 42

Cubillos, P. et al. 2013, The Astrophysical Journal, 768, 42

Dalcanton, J. et al. 2015, arXiv preprint arXiv:1507.04779

De Kok, R., Stam, D., & Karalidi, T. 2011, The Astrophysical Journal, 741, 59

De Maesschalck, R., Jouan-Rimbaud, D., & Massart, D. L. 2000, Chemometrics and intelligent laboratory systems, 50, 1

de Mooij, E. J., Brogi, M., de Kok, R., Snellen, I., Kenworthy, M., & Karjalainen, R. 2013, Astronomy & Astrophysics, 550, A54

de Wit, J., Gillon, M., Demory, B.-O., & Seager, S. 2012, A&A, 548, A128, 1202.3829

Deming, D., Harrington, J., Laughlin, G., Seager, S., Navarro, S. B., Bowman, W. C., & Horning, K. 2007, The Astrophysical Journal Letters, 667, L199

Deming, D., Harrington, J., Seager, S., & Richardson, L. J. 2006, The Astrophysical Journal, 644, 560

Deming, D. et al. 2011, The Astrophysical Journal, 726, 95

——. 2015, The Astrophysical Journal, 805, 132

Deming, D., Seager, S., Richardson, L. J., & Harrington, J. 2005, Nature, 434, 740

Deming, D. L., Berta-Thompson, Z., Cowan, N., Fortney, J. J., Kempton, E., Knutson, H., Rogers, L., & Sing, D. 2016, Report of the Hubble Space Telescope Exoplanet Committee

Demory, B.-O. et al. 2013, The Astrophysical Journal Letters, 776, L25

——. 2007, Astronomy & Astrophysics, 475, 1125

——. 2016a, Nature

Demory, B.-O., Gillon, M., Madhusudhan, N., & Queloz, D. 2016b, Monthly Notices of the Royal Astronomical Society, 455, 2018

Demory, B.-O., Gillon, M., Seager, S., Benneke, B., Deming, D., & Jackson, B. 2012, The Astrophysical Journal Letters, 751, L28

Demory, B.-O. et al. 2011, The Astrophysical Journal Letters, 735, L12

Désert, J.-M. et al. 2011a, The Astrophysical Journal Supplement Series, 197, 14

——. 2011b, The Astrophysical Journal Supplement Series, 197, 11

Désert, J.-M., Des Etangs, A. L., Hébrard, G., Sing, D. K., Ehrenreich, D., Ferlet, R., & Vidal-Madjar, A. 2009, The Astrophysical Journal, 699, 478

Désert, J.-M. et al. 2011c, Astronomy & Astrophysics, 526, A12

Diamond-Lowe, H., Stevenson, K. B., Bean, J. L., Line, M. R., & Fortney, J. J. 2014, The Astrophysical Journal, 796, 66, 1409.5336

Dlugach, J., & Yanovitskij, E. 1974, Icarus, 22, 66

Dones, L., & Tremaine, S. 1993, Icarus, 103, 67

Dong, S., & Zhu, Z. 2013, The Astrophysical Journal, 778, 53

Dressing, C. D., & Charbonneau, D. 2013, The Astrophysical Journal, 767, 95

Dunbar, B. 2009, NASA's Spitzer Telescope Warms Up To New Career

Ehrenreich, D., Hébrard, G., Des Etangs, A. L., Sing, D. K., Désert, J.-M., Bouchy, F., Ferlet, R., & Vidal-Madjar, A. 2007, The Astrophysical Journal Letters, 668, L179

Esteves, L. J., De Mooij, E. J., & Jayawardhana, R. 2014, arXiv preprint arXiv:1407.2245

Evans, T. et al. 2013, Astrophys. J. Lett, 772, L16

Evans, T. M., Aigrain, S., Gibson, N., Barstow, J. K., Amundsen, D. S., Tremblin, P., & Mourier, P. 2015, Monthly Notices of the Royal Astronomical Society, 451, 680

Fabrycky, D., & Tremaine, S. 2007, The Astrophysical Journal, 669, 1298

Faherty, J. K., Beletsky, Y., Burgasser, A. J., Tinney, C., Osip, D. J., Filippazzo, J. C., & Simcoe, R. A. 2014, The Astrophysical Journal, 790, 90

Fazio, G. et al. 2004, The Astrophysical Journal Supplement Series, 154, 10

Föhring, D., Dhillon, V., Madhusudhan, N., Marsh, T., Copperwheat, C., Littlefair, S., & Wilson, R. W. 2013, Monthly Notices of the Royal Astronomical Society, stt1443

Ford, E., Seager, S., & Turner, E. 2001, Nature, 412, 885

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, Publications of the Astronomical Society of the Pacific, 125, 306

Fortney, J., Marley, M., & Barnes, J. 2007, The Astrophysical Journal, 659, 1661

Fortney, J. J., Cooper, C. S., Showman, A. P., Marley, M. S., & Freedman, R. S. 2006, ApJ, 652, 746, astro-ph/0608235

Fortney, J. J. et al. 2011, The Astrophysical Journal Supplement Series, 197, 9

Fortney, J. J., Lodders, K., Marley, M. S., & Freedman, R. S. 2008, The Astrophysical Journal, 678, 1419

Fressin, F., Knutson, H. A., Charbonneau, D., ODonovan, F. T., Burrows, A., Deming, D., Mandushev, G., & Spiegel, D. 2010, The Astrophysical Journal, 711, 374

Fressin, F. et al. 2011, Nature

Fujii, Y., & Kawahara, H. 2012, The Astrophysical Journal, 755, 101

Fujii, Y., Kawahara, H., Suto, Y., Fukuda, S., Nakajima, T., Livengood, T. A., & Turner, E. L. 2011, The Astrophysical Journal, 738, 184

Fujii, Y., Kawahara, H., Suto, Y., Taruya, A., Fukuda, S., Nakajima, T., & Turner, E. L. 2010, The Astrophysical Journal, 715, 866

Gaidos, E., & Williams, D. 2004, New Astronomy, 10, 67

Gardner, J. P. et al. 2006, Space Science Reviews, 123, 485

Gautier III, T. N. et al. 2012, The Astrophysical Journal, 749, 15

Gibson, N., Aigrain, S., Barstow, J., Evans, T., Fletcher, L., & Irwin, P. 2013, Monthly Notices of the Royal Astronomical Society, 428, 3680

Gibson, N., Aigrain, S., Roberts, S., Evans, T., Osborne, M., & Pont, F. 2012, Monthly Notices of the Royal Astronomical Society, 419, 2683

Gibson, N. P., Aigrain, S., Barstow, J. K., Evans, T. M., Fletcher, L. N., & Irwin, P. G. J. 2013, MNRAS, 436, 2974, 1309.6998

Gibson, N. P., Pont, F., & Aigrain, S. 2011, Monthly Notices of the Royal Astronomical Society, 411, 2199

Gillon, M. et al. 2014, Astronomy & Astrophysics, 563, A21

——. 2016, Nature, 533, 221

——. 2010, Astronomy & Astrophysics, 511, A3

Gillon, M. et al. 2009, A&A, 506, 359

Ginzburg, S., Schlichting, H. E., & Sari, R. 2015, arXiv preprint arXiv:1512.07925

Goldblatt, C. 2016, arXiv preprint arXiv:1603.00950

Goldreich, P., & Soter, S. 1966, Icarus, 5, 375

Green, P. J. 1995, Biometrika, 82, 711

Grießmeier, J.-M. et al. 2004, Astronomy & Astrophysics, 425, 753

Grillmair, C., Charbonneau, D., Burrows, A., Armus, L., Stauffer, J., Meadows, V., Van Cleve, J., & Levine, D. 2007, The Astrophysical Journal Letters, 658, L115

Grillmair, C. J. et al. 2008, Nature, 456, 767

Grillmair, C. J. et al. 2012, in SPIE Astronomical Telescopes+ Instrumentation, International Society for Optics and Photonics, 84481I–84481I

Guillot, T. 2010, Astronomy & Astrophysics, 520, A27

Han, E., Wang, S. X., Wright, J. T., Feng, Y. K., Zhao, M., Fakhouri, O., Brown, J. I., & Hancock, C. 2014, Publications of the Astronomical Society of the Pacific, 126, 827

Hanel, R., Conrath, B., Herath, L., Kunde, V., & Pirraglia, J. 1981, Journal of Geophysical Research: Space Physics, 86, 8705

Hanel, R. A. 2003, Exploration of the solar system by infrared remote sensing (Cambridge University Press)

Hansen, C. J., Schwartz, J. C., & Cowan, N. B. 2014, Monthly Notices of the Royal Astronomical Society, 444, 3632

Harrington, J., Hansen, B. M., Luszcz, S. H., Seager, S., Deming, D., Menou, K., Cho, J. Y.-K., & Richardson, L. J. 2006, Science, 314, 623

Harrington, J., Luszcz, S., Seager, S., Deming, D., & Richardson, L. J. 2007, Nature, 447, 691

Haswell, C. A. 2010, Transiting exoplanets (Cambridge University Press)

Hebb, L. et al. 2009, The Astrophysical Journal, 693, 1920

Heller, R., Leconte, J., & Barnes, R. 2011, Astronomy & Astrophysics, 528, A27

Hellier, C. et al. 2011, Astronomy & Astrophysics, 535, L7

Heng, K., & Demory, B.-O. 2013, The Astrophysical Journal, 777, 100

Heng, K., Menou, K., & Phillipps, P. J. 2011, Monthly Notices of the Royal Astronomical Society, 413, 2380

Heng, K., & Showman, A. P. 2014, arXiv preprint arXiv:1407.4150

Holman, M. J., & Murray, N. W. 2005, Science, 307, 1288

Howell, S. B. et al. 2014, Publications of the Astronomical Society of the Pacific, 126, 398

Hu, R. 2014, arXiv preprint arXiv:1412.7582

Hubbard, W. B., Hattori, M., Burrows, A., Hubeny, I., & Sudarsky, D. 2007, Icarus, 187, 358

Ingalls, J. G., Carey, S. J., Lowrance, P. J., Grillmair, C. J., & Stauffer, J. R. 2014, in SPIE Astronomical Telescopes+ Instrumentation, International Society for Optics and Photonics, 91431M–91431M

Ingalls, J. G. et al. 2016, arXiv preprint arXiv:1601.05101

Jacob, D. 1999, Introduction to atmospheric chemistry (Princeton University Press)

Kataria, T., Showman, A. P., Fortney, J. J., Stevenson, K. B., Line, M. R., Kreidberg, L., Bean, J. L., & Désert, J.-M. 2015, The Astrophysical Journal, 801, 86

Kawahara, H. 2012, The Astrophysical Journal Letters, 760, L13

Kawahara, H., & Fujii, Y. 2010, The Astrophysical Journal, 720, 1333

——. 2011, The Astrophysical Journal Letters, 739, L62

Kimble, R. A., MacKenty, J. W., O'Connell, R. W., & Townsend, J. A. 2008, in SPIE Astronomical Telescopes+ Instrumentation, International Society for Optics and Photonics, 70101E–70101E

Kipping, D. M., & Spiegel, D. S. 2011, Monthly Notices of the Royal Astronomical Society: Letters, 417, L88

Kipping, D. M. et al. 2014, The Astrophysical Journal, 795, 25

Knutson, H. A., Charbonneau, D., Allen, L. E., Burrows, A., & Megeath, S. T. 2008, The Astrophysical Journal, 673, 526

Knutson, H. A. et al. 2007a, Nature, 447, 183

Knutson, H. A., Charbonneau, D., Burrows, A., O'Donovan, F. T., & Mandushev, G. 2009a, The Astrophysical Journal, 691, 866

Knutson, H. A., Charbonneau, D., Cowan, N. B., Fortney, J. J., Showman, A. P., Agol, E., & Henry, G. W. 2009b, The Astrophysical Journal, 703, 769

Knutson, H. A. et al. 2009c, The Astrophysical Journal, 690, 822

Knutson, H. A., Charbonneau, D., Deming, D., & Richardson, L. J. 2007b, Publications of the Astronomical Society of the Pacific, 119, 616

Knutson, H. A., Howard, A. W., & Isaacson, H. 2010, The Astrophysical Journal, 720, 1569

Knutson, H. A. et al. 2012, The Astrophysical Journal, 754, 22

Knutson, H. A. et al. 2011, ApJ, 735, 27, 1104.2901

Koch, D. G. et al. 2010, The Astrophysical Journal Letters, 713, L79

Kokubo, E., & Ida, S. 2007, The Astrophysical Journal, 671, 2082

Koll, D. D., & Abbot, D. S. 2015, The Astrophysical Journal, 802, 21

Konacki, M., Torres, G., Jha, S., & Sasselov, D. D. 2003, Nature, 421, 507

Koon, W. S., Lo, M. W., Marsden, J. E., & Ross, S. D. 2008, Free online Copy: Marsden Books

Kouveliotou, C. et al. 2014, arXiv preprint arXiv:1401.3741

Kreidberg, L. et al. 2014, The Astrophysical Journal Letters, 793, L27

Lagrange, A.-M. et al. 2009, Astronomy & Astrophysics, 493, L21

Lanotte, A. et al. 2014, Astronomy & Astrophysics, 572, A73

Laskar, J. 1994, Astronomy and Astrophysics, 287, L9

Laskar, J., Joutel, F., & Robutel, P. 1993, Nature, 361, 615

Laskar, J., Robutel, P., Joutel, F., Gastineau, M., Correia, A., & Levrard, B. 2004, Astronomy & Astrophysics, 428, 261

Latham, D. W., Mazeh, T., Stefanik, R. P., Mayor, M., & Burki, G. 1989

Lawson, P. et al. 2008, in SPIE Astronomical Telescopes + Instrumentation, International Society for Optics and Photonics, 70132N–70132N

Leary, J. C. et al. 2007, in The Messenger Mission to Mercury (Springer), 187–217

Lebreton, J.-P. et al. 2005, Nature, 438, 758

LeDrew, G. 2001, Journal of the Royal Astronomical Society of Canada, 95, 32

Lee, J.-M., Fletcher, L. N., & Irwin, P. G. 2012, Monthly Notices of the Royal Astronomical Society, 420, 170

Lewis, N. K. et al. 2013, The Astrophysical Journal, 766, 95

Li, S.-l., Miller, N., Lin, D. N., & Fortney, J. J. 2010, Nature, 463, 1054

Lin, D. N., Bodenheimer, P., & Richardson, D. C. 1996, Nature

Lindegren, L. et al. 2007, Proceedings of the International Astronomical Union, 3, 217

Line, M. R. et al. 2016, arXiv preprint arXiv:1605.08810

Line, M. R., Zhang, X., Vasisht, G., Natraj, V., Chen, P., & Yung, Y. L. 2012, The Astrophysical Journal, 749, 93

Lissauer, J. J. et al. 2012, The Astrophysical Journal, 750, 112

López-Morales, M., Coughlin, J. L., Sing, D. K., Burrows, A., Apai, D., Rogers, J. C., Spiegel, D. S., & Adams, E. R. 2010, The Astrophysical Journal Letters, 716, L36

Lyons, L. 1992, Journal of Physics A: Mathematical and General, 25, 1967

Machalek, P., Greene, T., McCullough, P. R., Burrows, A., Burke, C. J., Hora, J. L., Johns-Krull, C. M., & Deming, D. L. 2010, The Astrophysical Journal, 711, 111

Machalek, P., McCullough, P. R., Burke, C. J., Valenti, J. A., Burrows, A., & Hora, J. L. 2008, The Astrophysical Journal, 684, 1427

Machalek, P., McCullough, P. R., Burrows, A., Burke, C. J., Hora, J. L., & Johns-Krull, C. M. 2009, The Astrophysical Journal, 701, 514

Madhusudhan, N. 2012, The Astrophysical Journal, 758, 36

Madhusudhan, N. et al. 2011, Nature, 469, 64

Madhusudhan, N., & Seager, S. 2009, The Astrophysical Journal, 707, 24

——. 2010, The Astrophysical Journal, 725, 261

Mahtani, D. P. et al. 2013, MNRAS, 432, 693, 1303.4596

Majeau, C., Agol, E., & Cowan, N. B. 2012, ApJL, 747, L20, 1202.1883

Mandell, A. M., Deming, L. D., Blake, G. A., Knutson, H. A., Mumma, M. J., Villanueva, G. L., & Salyk, C. 2011, The Astrophysical Journal, 728, 18

Marley, M. S., Gelino, C., Stephens, D., Lunine, J. I., & Freedman, R. 1999, The Astrophysical Journal, 513, 879

Marley, M. S., Saumon, D., Cushing, M., Ackerman, A. S., Fortney, J. J., & Freedman, R. 2012, The Astrophysical Journal, 754, 135

Marois, C., Macintosh, B., Barman, T., Zuckerman, B., Song, I., Patience, J., Lafrenière, D., & Doyon, R. 2008, Science, 322, 1348

Martin, T., Leconte, J., Selsis, F., Emeline, B., Francois, F., Ignasi, R., Raymond, S. N., & Guillem, A.-E. 2016, Astronomy and Astrophysics-A&A, arXiv

Matson, D. L., Spilker, L. J., & Lebreton, J.-P. 2003, in The Cassini-Huygens Mission (Springer), 1–58

Maxted, P. et al. 2013, Monthly Notices of the Royal Astronomical Society, 428, 2645

Mayor, M., & Queloz, D. 1995, A Jupiter-mass companion to a solar-type star

Meadows, V. S. et al. 2016, arXiv preprint arXiv:1608.08620

Miguel, Y., & Brunini, A. 2010, Monthly Notices of the Royal Astronomical Society, 406, 1935

Miralda-Escudé, J. 2002, The Astrophysical Journal, 564, 1019

Mitsuda, K. et al. 1984, PASJ, 36, 741

Morello, G., Waldmann, I. P., Tinetti, G., Peres, G., Micela, G., & Howarth, I. D. 2014, The Astrophysical Journal, 786, 22

Morton, T. D., Bryson, S. T., Coughlin, J. L., Rowe, J. F., Ravichandran, G., Petigura, E. A., Haas, M. R., & Batalha, N. M. 2016, The Astrophysical Journal, 822, 86

Motalebi, F. et al. 2015, Astronomy & Astrophysics, 584, A72

Mullally, F. et al. 2015, The Astrophysical Journal Supplement Series, 217, 31

Nesvornỳ, D., Kipping, D. M., Buchhave, L. A., Bakos, G. Á., Hartman, J., & Schmitt, A. R. 2012, Science, 336, 1133

Nikolov, N., & Sainsbury-Martinez, F. 2015, The Astrophysical Journal, 808, 57

Nymeyer, S. et al. 2011, The Astrophysical Journal, 742, 35

Oakley, P., & Cash, W. 2009, The Astrophysical Journal, 700, 1428

O'Donovan, F. T., Charbonneau, D., Harrington, J., Madhusudhan, N., Seager, S., Deming, D., & Knutson, H. A. 2010, The Astrophysical Journal, 710, 1551

O'Rourke, J. G. et al. 2014, The Astrophysical Journal, 781, 109

Pallé, E., Ford, E. B., Seager, S., Montañés-Rodríguez, P., & Vazquez, M. 2008, The Astrophysical Journal, 676, 1319

Papaloizou, J., Nelson, R., Kley, W., Masset, F., & Artymowicz, P. 2007, Protostars and Planets V, 1, 655

Papaloizou, J. C., & Terquem, C. 2005, Reports on Progress in Physics, 69, 119

Perez-Becker, D., & Showman, A. P. 2013, The Astrophysical Journal, 776, 134

Petigura, E. A., Howard, A. W., & Marcy, G. W. 2013, Proceedings of the National Academy of Sciences, 110, 19273

Pierrehumbert, R. T. 2010, Principles of planetary climate (Cambridge University Press)

Pollack, J. B., Podolak, M., Bodenheimer, P., & Christofferson, B. 1986, Icarus, 67, 409

Pollack, J. B., Toon, O. B., & Boese, R. 1980, Journal of Geophysical Research: Space Physics, 85, 8223

Pont, F., Sing, D., Gibson, N., Aigrain, S., Henry, G., & Husnoo, N. 2013, Monthly Notices of the Royal Astronomical Society, stt651

Pont, F., Zucker, S., & Queloz, D. 2006, Monthly Notices of the Royal Astronomical Society, 373, 231

Postman, M. et al. 2010, in Pathways Towards Habitable Planets, Vol. 430, 361

Queloz, D. et al. 2010, Astronomy & Astrophysics, 517, L1

Qui, J. et al. 2003, Journal of Geophysical Research, 12

Ranjan, S., Charbonneau, D., Désert, J.-M., Madhusudhan, N., Deming, D., Wilkins, A., & Mandell, A. M. 2014, The Astrophysical Journal, 785, 148

Rasio, F. A., & Ford, E. B. 1996, Science, 274, 954

Rauscher, E., & Kempton, E. M. R. 2014, ApJ, 790, 79, 1402.4833

Rauscher, E., Menou, K., Cho, J. Y.-K., Seager, S., & Hansen, B. M. 2007, The Astrophysical Journal Letters, 662, L115

Reach, W. T. et al. 2005, Publications of the Astronomical Society of the Pacific, 117, 978

Ribas, I. et al. 2016, arXiv preprint arXiv:1608.06813

Richardson, L. J., Deming, D., & Seager, S. 2003, The Astrophysical Journal, 597, 581

Ricker, G. R. et al. 2015, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003

Robinson, T. D., Ennico, K., Meadows, V. S., Sparks, W., Bussey, D. B. J., Schwieterman, E. W., & Breiner, J. 2014, The Astrophysical Journal, 787, 171

Robinson, T. D., Meadows, V. S., & Crisp, D. 2010, The Astrophysical Journal Letters, 721, L67

Rogers, J., López-Morales, M., Apai, D., & Adams, E. 2013, The Astrophysical Journal, 767, 64

Rogers, J. C., Mercedes, L., Sing, D. K., Burrows, A., et al. 2009, The Astrophysical Journal, 707, 1707

Rogers, L. A. 2015, The Astrophysical Journal, 801, 41

Rostron, J. W., Wheatley, P. J., Anderson, D. R., Cameron, A. C., Fortney, J. J., Harrington, J., Knutson, H. A., & Pollacco, D. L. 2014, Monthly Notices of the Royal Astronomical Society, 441, 3666

Rowe, J. F. et al. 2014, The Astrophysical Journal, 784, 45

Rowe, J. F. et al. 2006, ApJ, 646, 1241, astro-ph/0603410

Rowe, J. F. et al. 2008, The Astrophysical Journal, 689, 1345

Roy, D. P. et al. 2014, Remote Sensing of Environment, 145, 154

Schlichting, H. E., & Sari, R. 2007, The Astrophysical Journal, 658, 593

Schneider, J., Dedieu, C., Le Sidaner, P., Savalle, R., & Zolotukhin, I. 2011, Astronomy & Astrophysics, 532, A79

Schwartz, J. C., & Cowan, N. B. 2015, Monthly Notices of the Royal Astronomical Society, 449, 4192

Schwartz, J. C., Sekowski, C., Haggard, H. M., Pallé, E., & Cowan, N. B. 2016, Monthly Notices of the Royal Astronomical Society, 457, 926

Schwarz, G., et al. 1978, The annals of statistics, 6, 461

Schwarz, H., Brogi, M., de Kok, R., Birkby, J., & Snellen, I. 2015, ArXiv e-prints, 1502.04713

Seager, S. 2013, Science, 340, 577

Seager, S., & Deming, D. 2010, arXiv preprint arXiv:1005.4037

Seager, S., & Hui, L. 2002, The Astrophysical Journal, 574, 1004

Seager, S., Kuchner, M., Hier-Majumder, C., & Militzer, B. 2007, The Astrophysical Journal, 669, 1279

Seager, S., Richardson, L., Hansen, B., Menou, K., Cho, J.-K., & Deming, D. 2005, The Astrophysical Journal, 632, 1122

Shabram, M., Fortney, J. J., Greene, T. P., & Freedman, R. S. 2011, The Astrophysical Journal, 727, 65

Showman, A. P., Cho, J. Y., & Menou, K. 2010, Exoplanets, 526, 471

Showman, A. P., Lewis, N. K., & Fortney, J. J. 2014, ArXiv e-prints, 1411.4731

Shporer, A. et al. 2014, The Astrophysical Journal, 788, 92

Silburt, A., Gaidos, E., & Wu, Y. 2015, The Astrophysical Journal, 799, 180

Sing, D. K. et al. 2011, Monthly Notices of the Royal Astronomical Society, 416, 1443

Singh, S. 2010, Serendipity: Fortune and the Prepared Mind, 22, 65

Skilling, J. 2004, Bayesian inference and maximum entropy methods in science and engineering, 735, 395

Smith, A. et al. 2012, Astronomy & Astrophysics, 545, A93

Spiegel, D. S., Menou, K., & Scharf, C. A. 2009, The Astrophysical Journal, 691, 596

Stern, S. et al. 2015, Science, 350, aad1815

Stevenson, K. B., Bean, J. L., Madhusudhan, N., & Harrington, J. 2014a, The Astrophysical Journal, 791, 36

Stevenson, K. B., Bean, J. L., Seifahrt, A., Désert, J.-M., Madhusudhan, N., Bergmann, M., Kreidberg, L., & Homeier, D. 2014b, The Astronomical Journal, 147, 161

Stevenson, K. B., Bean, J. L., Seifahrt, A., Gilbert, G. J., Line, M. R., Désert, J.-M., & Fortney, J. J. 2016a, The Astrophysical Journal, 817, 141

Stevenson, K. B. et al. 2014c, Science, 346, 838

——. 2012a, The Astrophysical Journal, 754, 136

——. 2012b, The Astrophysical Journal, 755, 9

——. 2010, Nature, 464, 1161

——. 2016b, arXiv preprint arXiv:1608.00056

Stone, E., & Miner, E. 1986, Science, 233, 39

——. 1989, Science, 246, 1417

Sudarsky, D., Burrows, A., & Pinto, P. 2000, The Astrophysical Journal, 538, 885

Swain, M. et al. 2013, Icarus, 225, 432

Swain, M. G. et al. 2010, Gastroenterology, 139, 1593

Swain, M. R., Vasisht, G., & Tinetti, G. 2008, Nature, 452, 329

Szabó, G. M. et al. 2011, The Astrophysical Journal Letters, 736, L4

Team, L. S. P. E. 1969, Science, 165, 1211

Tinetti, G., Deroo, P., Swain, M., Griffith, C., Vasisht, G., Brown, L., Burke, C., & McCullough, P. 2010, The Astrophysical Journal Letters, 712, L139

Tinetti, G. et al. 2007, Nature, 448, 169

Tinsley, B. M. 1980, Fundamentals of cosmic physics, 5, 287

Todorov, K., Deming, D., Harrington, J., Stevenson, K. B., Bowman, W. C., Nymeyer, S., Fortney, J. J., & Bakos, G. A. 2009, The Astrophysical Journal, 708, 498

Todorov, K. O., Deming, D., Burrows, A., & Grillmair, C. J. 2014, The Astrophysical Journal, 796, 100

Todorov, K. O. et al. 2013, The Astrophysical Journal, 770, 102

——. 2012, The Astrophysical Journal, 746, 111

Tong, Y. L. 2012, The multivariate normal distribution (Springer Science & Business Media)

Topping, J., & Worrell, F. T. 1957, American Journal of Physics, 25, 498

Trauger, J. T., & Traub, W. A. 2007, Nature, 446, 771

Triaud, A. H. 2014, Monthly Notices of the Royal Astronomical Society: Letters, 439, L61

Triaud, A. H. et al. 2015, Monthly Notices of the Royal Astronomical Society, 450, 2279

Turnbull, M. C. et al. 2012, Publications of the Astronomical Society of the Pacific, 124, 418

Unwin, S. C. et al. 2008, Publications of the Astronomical Society of the Pacific, 120, 38

Valencia, D., O'Connell, R. J., & Sasselov, D. 2006, Icarus, 181, 545

Valencia, D., Sasselov, D. D., & O'Connell, R. J. 2007a, The Astrophysical Journal, 656, 545

Valencia, D., Sasselov, D. D., & OConnell, R. J. 2007b, The Astrophysical Journal, 665, 1413

Van Dyk, S., Werner, M., & Silbermann, N. 2013, Spitzer Space Telescope Handbook

Waldmann, I. P. 2012, The Astrophysical Journal, 747, 12

Waldmann, I. P., Tinetti, G., Deroo, P., Hollis, M. D., Yurchenko, S. N., & Tennyson, J. 2013, The Astrophysical Journal, 766, 7

Wall, M. 2015, How Will the Hubble Space Telescope Die?

Wang, W., van Boekel, R., Madhusudhan, N., Chen, G., Zhao, G., & Henning, T. 2013, The Astrophysical Journal, 770, 70

Weiss, L. M., & Marcy, G. W. 2014, The Astrophysical Journal Letters, 783, L6

Werner, M. et al. 2004, The Astrophysical Journal Supplement Series, 154, 1

Wheatley, P. J. et al. 2010, arXiv preprint arXiv:1004.0836

Wilkins, A. N., Deming, D., Madhusudhan, N., Burrows, A., Knutson, H., McCullough, P., & Ranjan, S. 2014, The Astrophysical Journal, 783, 113

Williams, D. M., & Gaidos, E. 2008, Icarus, 195, 927

Williams, D. M., & Kasting, J. F. 1997, Icarus, 129, 254

Williams, D. M., & Pollard, D. 2003, International Journal of Astrobiology, 2, 1

Winker, D. M., Pelon, J. R., & McCormick, M. P. 2003, in Third International Asia-Pacific Environmental Remote Sensing Remote Sensing of the Atmosphere, Ocean, Environment, and Space, International Society for Optics and Photonics, 1–11

Wolszczan, A., et al. 1994, Science-AAAS-Weekly Paper Edition-including Guide to Scientific Information, 264, 538

Wong, I. et al. 2014, The Astrophysical Journal, 794, 134

——. 2016, The Astrophysical Journal, 823, 122

——. 2015, The Astrophysical Journal, 811, 122

Wright, J., & Gaudi, B. 2013, TD Oswalt, LM French, & P. Kalas, 489

Wright, J., Marcy, G., Howard, A., Johnson, J. A., Morton, T., & Fischer, D. 2012, The Astrophysical Journal, 753, 160

Wright, J. T. et al. 2011, Publications of the Astronomical Society of the Pacific, 123, 412

Xiong, X., Chiang, K., Sun, J., Barnes, W., Guenther, B., & Salomonson, V. 2009, Advances in Space Research, 43, 413

Yang, J., Cowan, N. B., & Abbot, D. S. 2013, The Astrophysical Journal Letters, 771, L45

Yu, L. et al. 2015, The Astrophysical Journal, 812, 48

Zahnle, K., Marley, M. S., Freedman, R., Lodders, K., & Fortney, J. 2009, The Astrophysical Journal Letters, 701, L20

Zellem, R. T. et al. 2014, The Astrophysical Journal, 790, 53

Zhao, M., Monnier, J. D., Swain, M. R., Barman, T., & Hinkley, S. 2012, The Astrophysical Journal, 744, 122

Zhao, M. et al. 2014, ApJ, 796, 115, 1410.0968

Zhou, G., Bayliss, D., Kedziora-Chudczer, L., Salter, G., Tinney, C., & Bailey, J. 2014, Monthly Notices of the Royal Astronomical Society, 445, 2746

Zhou, G., Bayliss, D., Kedziora-Chudczer, L., Tinney, C., Bailey, J., Salter, G., & Rodriguez, J. 2015, Monthly Notices of the Royal Astronomical Society, 454, 3002

Zhou, G., Kedziora-Chudczer, L., Bayliss, D. D., & Bailey, J. 2013, The Astrophysical Journal, 774, 118

APPENDIX A

BLISS: Choosing Parameters

A.1. Coefficients for the Pixel Sensitivity

To pick the $c_{\ell m}$ in Equation 3.3, we start by choosing how much all the terms added together can change the sensitivity, or a_v. We divide this value by how many coefficients we have, n_c, where we use $n_c = 35$ because we set $n = 7$. We cannot give each polynomial term the same magnitude everywhere on the pixel, so we scale the terms to be the same at some reference distance, d_{ref}, from the pixel center. This gives us the equation:

$$\mathbb{C}_{\ell m} = \frac{a_v}{n_c} \left(\frac{1}{d_{\mathrm{ref}}} \right)^{\ell + m}, \tag{A.1}$$

where $\mathbb{C}_{\ell m}$ is a limit for each coefficient. By doing this, the lower-order terms will dominate inside d_{ref} and vice versa, so the sensitivity tends to vary more near the pixel edges. Note that the pixel centers have the highest sensitivities in the real IRAC detector (e.g. Reach et al., 2005; Cowan et al., 2012b), which is not always true in this model. We decide to set $a_v = 0.5$ and $d_{\mathrm{ref}} = 0.1$, but other choices work, too.

We next randomly pick each $c_{\ell m} \in [-\mathbb{C}_{\ell m}, \mathbb{C}_{\ell m}]$, then *rescale* all these coefficients to get a chosen amplitude for the detector signal, ΔD, no matter what centroids we have. For each new sensitivity map, we draw and rescale the $c_{\ell m}$ again.

A.2. Eclipse and Phase Curve

To choose the parameters for Equations 3.5 and 3.6, we start with the eclipse and work backwards. We fix $t_{\max} = 6$ hrs and $t_w = 1$ hr, and because we randomly choose $t_e \in [2, 4]$ hrs, there is always some baseline before and after the eclipse. Then we pick a value of ΔD_e—we set δ_e first (to 5.0×10^{-3}) if we are fitting the light curve via MCMC (Section 3.4.3) and ΔD first if making a BLISS map (Section 3.4.2). In the second case, the eclipse depth is about 10^{-5}–10^{-2}.

Then we look at the phase model. We randomly pick $P_{\text{orb}} \in [15, 60]$ hrs, which gives us part of a phase curve, and $\phi_o \in \left[\pi\left(1 - \frac{12}{P_{\text{orb}}}\right), \pi\right]$, which makes sure the peak of the phase curve happens during the observation. The bottom of the eclipse should be lower in flux than the phase curve *could* be, so we calculate the maximum half-amplitude, $\alpha_{\max}$, the phase curve could have given the other parameters. Then we randomly choose $\alpha \in [0.7\alpha_{\max}, \alpha_{\max}]$, where the lower limit on α could be different and is just by choice.

Lastly, we pick the amount of photon noise, σ, depending on how significant we want the eclipse to be (Equation 3.8). Our choices give us light curves that mimic real data; other choices could work as well.

APPENDIX B

Obliquity: Describing Planets and Kernels

B.1. Viewing Geometry

B.1.1. General Observer

The time-dependence of the kernel is contained in the sub-observer and sub-stellar angles: θ_o, ϕ_o θ_s, ϕ_s. Since they do not depend on planetary latitude or longitude, these four angles may be factored out of the kernel integrals. However, the light curves are still functions of time, so we derive the relevant dependencies here.

In particular, we compute the sub-stellar and sub-observer locations for planets on circular orbits using seven parameters. Three are intrinsic to the system: rotational angular frequency, $\omega_{\rm rot} \in (-\infty, \infty)$, orbital angular frequency, $\omega_{\rm orb} \in (0, \infty)$, and obliquity, $\Theta \in [0, \pi/2]$. Rotational frequency is measured in an inertial frame, where positive values are prograde and negative denotes retrograde rotation (for comparison, the rotational frequency of Earth is $\omega_{\rm rot}^{\oplus} \approx 2\pi/23.93$ h^{-1}). Two more parameters are extrinsic and differ for each observer: orbital inclination, $i \in [0, \pi/2]$ where $i = 90°$ is edge-on, and solstice phase, $\xi_{\rm s} \in [0, 2\pi)$, which is the orbital angle between superior conjunction and the maximum Northern excursion of the sub-stellar point. The remaining parameters are extrinsic initial conditions: the starting orbital position, $\xi_0 \in [0, 2\pi)$, and the initial sub-observer longitude, $\phi_o(0) \in [0, 2\pi)$. These parameters are illustrated in Figure B.1; other combinations are possible.

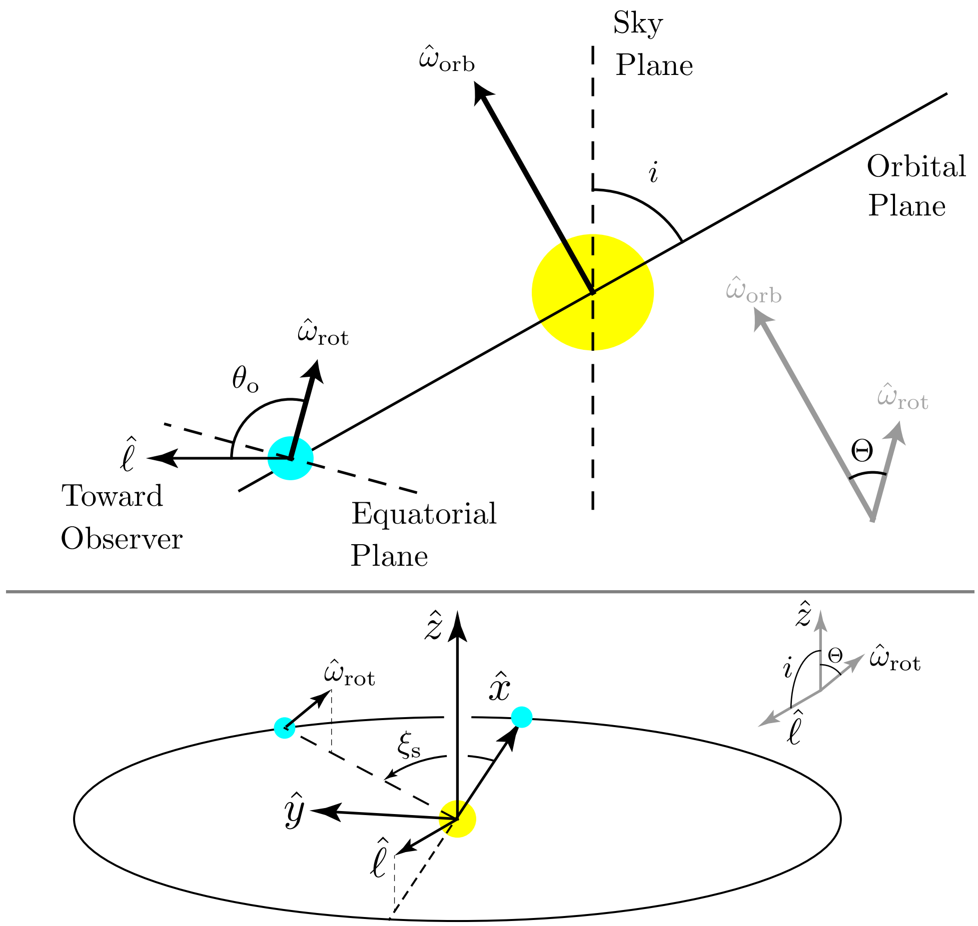


Figure B.1. The upper panel shows a side view of the general planetary system. The rotational and orbital angular frequencies $\{\omega_{\rm rot}, \omega_{\rm orb}\}$, inclination i, obliquity Θ, sub-observer colatitude θ_o, and observer viewing direction $\hat{\ell}$ are indicated. The lower panel is an isometric view, showing the solstice phase $\xi_{\rm s}$. Superior conjunction occurs along the positive x-axis. Note the inertial coordinates, how they relate to the observer's viewpoint and planet's spin axis, and the angles between these vectors.

We define the orbital phase of the planet as $\xi(t) = \omega_{\rm orb}t + \xi_0$. Without loss of generality we may set the first initial condition as $\xi_0 = 0$, which puts the planet at superior conjunction when $t = 0$. With no precession the sub-observer colatitude is constant,

$$\theta_o(t) = \theta_o. \tag{B.1}$$

This angle can be expressed in terms of the inclination, obliquity, and solstice phase using the spherical law of cosines (bottom of Figure B.1):

$$\cos\theta_o = \cos i \cos\Theta + \sin i \sin\Theta \cos\xi_{\mathrm{s}}, \tag{B.2}$$

$$\sin\theta_o = \sqrt{1 - \cos^2\theta_o}. \tag{B.3}$$

The sub-observer longitude decreases linearly with time for prograde rotation, as we define longitude increasing to the East:

$$\phi_o(t) = -\omega_{\mathrm{rot}} t + \phi_o(0). \tag{B.4}$$

The prime meridian ($\phi_p \Rightarrow \phi = 0$) is a free parameter, which we define to run from the planet's North pole to the sub-observer point at $t = 0$. This sets the second initial condition, namely $\phi_o(0) = 0$, and means

$$\cos\phi_o = \cos\left(-\omega_{\mathrm{rot}} t\right), \tag{B.5}$$

$$\sin\phi_o = \sqrt{1 - \cos^2\phi_o}. \tag{B.6}$$

Hence, the time evolution of the sub-observer point is specified by its colatitude and the rotational angular frequency.

The sub-stellar position is more complex for planets with non-zero obliquity. Consider an inertial Cartesian frame centered on the host star with fixed axes as follows: the z-axis is along the orbital angular frequency, $\hat{z} = \hat{\omega}_{\rm orb}$, while the x-axis points towards superior conjunction. The y-axis is then orthogonal to this plane using $\hat{y} = \hat{z} \times \hat{x}$ (bottom of Figure B.1). In these inertial coordinates, the unit vector from the planet center towards the host star is $\hat{r}_{\rm ps} = -\cos\xi\hat{x} - \sin\xi\hat{y}$. The corresponding unit vector from the star towards the observer is $\hat{\ell} = -\sin i\hat{x} + \cos i\hat{z}$. Our approach is to express everything in the inertial coordinate system, then find the sub-stellar point with appropriate dot products.

For the planetary surface, we use a second coordinate system fixed to the planet. We align the z_p-axis with the rotational angular frequency, $\hat{z}_p = \hat{\omega}_{\rm rot}$, while the x_p-axis is set by our choice for the prime meridian (and initial sub-observer longitude.) The final axis, y_p, is again determined by taking $\hat{y}_p = \hat{z}_p \times \hat{x}_p$. We proceed in two steps, first finding the planetary axes when $t = 0$, then using the planet's rotation to describe these axes at any time.

Since we disregard precession, the planet's rotation axis is time-independent:

$$\hat{z}_p = \hat{\omega}_{\rm rot} = -\cos\xi_{\rm s}\sin\Theta\hat{x} - \sin\xi_{\rm s}\sin\Theta\hat{y} + \cos\Theta\hat{z}. \tag{B.7}$$

The sub-observer point is on the prime meridian when $t = 0$, so that

$$\hat{y}_p(0) = \frac{\hat{z}_p \times \hat{\ell}}{\sin\theta_o}. \tag{B.8}$$

Computing this we find

$$\begin{aligned}\hat{y}_p(0) = \frac{1}{\sin\theta_o}\Big[&-\cos i \sin\xi_{\mathrm{s}} \sin\Theta\hat{x} \\ &+ (\cos i \cos\xi_{\mathrm{s}} \sin\Theta - \sin i \cos\Theta)\hat{y} \\ &- \sin i \sin\xi_{\mathrm{s}} \sin\Theta\hat{z}\Big].\end{aligned} \tag{B.9}$$

The starting x_p-axis is then found by taking $\hat{y}_p(0) \times \hat{z}_p$. The result is simplified by using Equation B.2:

$$\begin{aligned}\hat{x}_p(0) = \frac{1}{\sin\theta_o}\Big[&(\cos\xi_{\mathrm{s}} \sin\Theta \cos\theta_o - \sin i)\hat{x} \\ &+ \sin\xi_{\mathrm{s}} \sin\Theta \cos\theta_o\hat{y} \\ &+ \sin\Theta(\cos i \sin\Theta - \sin i \cos\xi_{\mathrm{s}} \cos\Theta)\hat{z}\Big].\end{aligned} \tag{B.10}$$

We can now find the planetary axes, in terms of the inertial axes, at any time by rotating Equations B.9 and B.10 about the z_p-axis:

$$\hat{x}_p = \cos(\omega_{\mathrm{rot}}t)\hat{x}_p(0) + \sin(\omega_{\mathrm{rot}}t)\hat{y}_p(0), \tag{B.11}$$

$$\hat{y}_p = -\sin(\omega_{\mathrm{rot}}t)\hat{x}_p(0) + \cos(\omega_{\mathrm{rot}}t)\hat{y}_p(0). \tag{B.12}$$

The sub-stellar angles in the planetary coordinates may then be extracted from the relations

$$\sin\theta_s \cos\phi_s = \hat{r}_{\mathrm{ps}} \cdot \hat{x}_p, \tag{B.13}$$

(B.14) $$\sin\theta_s \sin\phi_s = \hat{r}_{\mathrm{ps}} \cdot \hat{y}_p,$$

(B.15) $$\cos\theta_s = \hat{r}_{\mathrm{ps}} \cdot \hat{z}_p,$$

resulting in

(B.16) $$\cos\theta_s = \sin\Theta \cos\left[\xi - \xi_{\mathrm{s}}\right],$$

(B.17) $$\sin\theta_s = \sqrt{1 - \sin^2\Theta \cos^2\left[\xi - \xi_{\mathrm{s}}\right]},$$

(B.18) $$\cos\phi_s = \frac{\cos(\omega_{\mathrm{rot}} t) a(t) + \sin(\omega_{\mathrm{rot}} t) b(t)}{\sqrt{1 - \cos^2\theta_o}\sqrt{1 - \sin^2\Theta \cos^2\left[\xi - \xi_{\mathrm{s}}\right]}},$$

(B.19) $$\sin\phi_s = \frac{-\sin(\omega_{\mathrm{rot}} t) a(t) + \cos(\omega_{\mathrm{rot}} t) b(t)}{\sqrt{1 - \cos^2\theta_o}\sqrt{1 - \sin^2\Theta \cos^2\left[\xi - \xi_{\mathrm{s}}\right]}},$$

where the factors $a(t)$ and $b(t)$ are given by

(B.20) $$a(t) = \Big\{ \sin i \cos\xi - \cos\theta_o \sin\Theta \cos\left[\xi - \xi_{\mathrm{s}}\right] \Big\},$$

(B.21) $$b(t) = \Big\{ \sin i \sin\xi \cos\Theta - \cos i \sin\Theta \sin\left[\xi - \xi_{\mathrm{s}}\right] \Big\}.$$

Note that when $\theta_s = \{0, \pi\}$, the sub-stellar longitude can be set arbitrarily to avoid dividing by zero in Equations B.18 and B.19.

B.1.2. Polar Observer

Equations B.18 and B.19 for the sub-stellar longitude apply to most observers. However, the definition of $\hat{y}_p(0)$ in Equation B.8 fails when the sub-observer point coincides with one of the planet's poles. Two alternate definitions can be used in these situations.

Case 1: If the sub-stellar point *will not* pass over the poles during orbit, we may define

$$\hat{y}_p(0) = -\hat{y}, \tag{B.22}$$

so that

$$\hat{x}_p(0) = \hat{y}_p(0) \times \hat{z}_p = -\cos\Theta\hat{x} - \cos\xi_{\mathrm{s}}\sin\Theta\hat{z}. \tag{B.23}$$

This results in

$$\cos\phi_s = \frac{\cos\omega_{\mathrm{rot}}\cos\xi\cos\Theta + \sin\omega_{\mathrm{rot}}\sin\xi}{\sqrt{1-\sin^2\Theta\cos^2\left[\xi-\xi_{\mathrm{s}}\right]}}, \tag{B.24}$$

$$\sin\phi_s = \frac{-\sin\omega_{\mathrm{rot}}\cos\xi\cos\Theta + \cos\omega_{\mathrm{rot}}\sin\xi}{\sqrt{1-\sin^2\Theta\cos^2\left[\xi-\xi_{\mathrm{s}}\right]}}. \tag{B.25}$$

Case 2: However, if the sub-stellar point *will* pass over the poles during orbit, we define instead

$$\hat{x}_p(0) = \hat{z}, \tag{B.26}$$

such that

$$\hat{y}_p(0) = \hat{z}_p \times \hat{x}_p(0) = \cos\xi_{\mathrm{s}}\hat{y}. \tag{B.27}$$

This produces

$$\cos\phi_s = \frac{-\sin\omega_{\mathrm{rot}}\sin\xi\cos\xi_{\mathrm{s}}}{\sqrt{1-\sin^2\Theta\cos^2[\xi-\xi_{\mathrm{s}}]}}, \tag{B.28}$$

$$\sin\phi_s = \frac{-\cos\omega_{\mathrm{rot}}\sin\xi\cos\xi_{\mathrm{s}}}{\sqrt{1-\sin^2\Theta\cos^2[\xi-\xi_{\mathrm{s}}]}}. \tag{B.29}$$

These special cases *only* impact the sub-stellar longitude: expressions for the other angles are unchanged. As with a general observer, the Case 2 sub-stellar longitude may be set arbitrarily whenever $\theta_s = \{0, \pi\}$.

B.1.3. Zero Obliquity

For non-oblique planets, $\Theta = 0°$, the sub-observer colatitude satisfies

$$\cos\theta_o = \cos i, \tag{B.30}$$

$$\sin\theta_o = \sin i, \tag{B.31}$$

while the sub-stellar angles become

$$\cos\theta_s = 0, \tag{B.32}$$

$$\sin\theta_s = 1, \tag{B.33}$$

$$\cos\phi_s = \frac{\cos(\omega_{\rm rot}t)c(t) + \sin(\omega_{\rm rot}t)d(t)}{\sqrt{1-\cos^2 i}}, \tag{B.34}$$

$$\sin\phi_s = \frac{-\sin(\omega_{\rm rot}t)c(t) + \cos(\omega_{\rm rot}t)d(t)}{\sqrt{1-\cos^2 i}}, \tag{B.35}$$

where $c(t)$ and $d(t)$ are given by

$$c(t) = \sin i\cos\xi, \tag{B.36}$$

$$d(t) = \sin i\sin\xi. \tag{B.37}$$

The sub-stellar longitude is therefore

$$\begin{aligned}\cos\phi_s &= \frac{\cos(\omega_{\rm rot}t)\sin i\cos\xi + \sin(\omega_{\rm rot}t)\sin i\sin\xi}{\sqrt{1-\cos^2 i}} \\ &= \cos(\omega_{\rm rot}t)\cos\xi + \sin(\omega_{\rm rot}t)\sin\xi \\ &= \cos(\xi - \omega_{\rm rot}t),\end{aligned} \tag{B.38}$$

$$\begin{aligned}\sin\phi_s &= \frac{-\sin(\omega_{\rm rot}t)\sin i\cos\xi + \cos(\omega_{\rm rot}t)\sin i\sin\xi}{\sqrt{1-\cos^2 i}} \\ &= -\sin(\omega_{\rm rot}t)\cos\xi + \cos(\omega_{\rm rot}t)\sin\xi \\ &= \quad \sin(\xi - \omega_{\rm rot}t).\end{aligned} \tag{B.39}$$

In other words, $\theta_o = i$, $\phi_o = \phi_o(0) - \omega_{\rm rot}t$, $\theta_s = 0$, and $\phi_s = \xi - \omega_{\rm rot}t$.

B.2. Kernel Details

B.2.1. Characteristics

An important measure of the longitudinal kernel is its width, σ_ϕ, as shown in the left panel of Figure B.2. We treat this width mathematically as a standard deviation. Since $K(\phi, \mathbb{G})$ is on a periodic domain, we minimize the variance for each geometry with respect to the grid location of the prime meridian, ϕ_p:

$$\sigma_\phi^2 = \min \left[\int_0^{2\pi} \left(\phi^{'} - \bar{\phi} \right)^2 \hat{K}(\phi) \mathrm{d}\phi \right]_{\phi_p}, \tag{B.40}$$

where $\hat{K}(\phi) = K(\phi) / \int K(\phi)\mathrm{d}\phi$ is the spherically normalized longitudinal kernel, $\phi^{'} \equiv \phi + \phi_p$, and $\bar{\phi}$ is the mean longitude:

$$\bar{\phi} = \int_0^{2\pi} \phi^{'} \hat{K}(\phi) \mathrm{d}\phi. \tag{B.41}$$

All longitude arguments and separations in Equations B.40 and B.41 wrap around the standard domain $[0, 2\pi)$. Also note the *unprimed* arguments inside the kernel: these make computing the variance simpler. The minimum variance determines the standard deviation of the kernel, and thus width, for a given geometry.

The dominant colatitude is similarly important for the latitudinal kernel, as shown in the right panel of Figure B.2. Cowan et al. (2012a) defined the dominant colatitude, $\bar{\theta}$, for a given geometry:

$$\bar{\theta} = \oint \theta \hat{K}(\theta, \phi) \mathrm{d}\Omega, \tag{B.42}$$

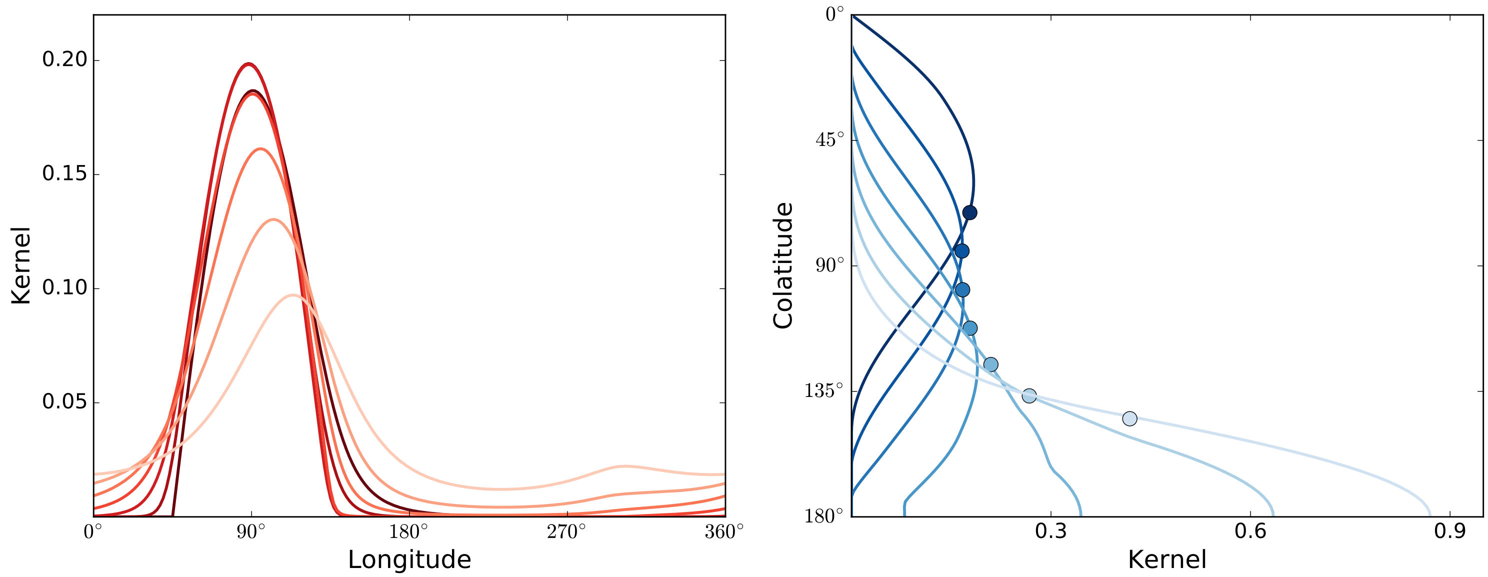


Figure B.2. *Left:* Longitudinal kernels of a planet with different obliquities at first quarter phase, $\xi(t) = 90°$, where $\xi_s = 225°$ and $i = 60°$. The curves show 0°–90° obliquity in 15° increments, where lighter shades of red indicate higher axial tilts. The kernel width of this planet changes as obliquity increases. *Right:* Analogous latitudinal kernels that indicate the dominant colatitude, shown as a circle, also changes as obliquity increases.

where $\hat{K}(\theta,\phi) = K(\theta,\phi)/\oint K(\theta,\phi)\mathrm{d}\Omega$ is the normalized kernel. Equation B.42 is equivalent to

$$\bar{\theta} = \int_0^{\pi} \theta\hat{K}(\theta)\sin\theta\mathrm{d}\theta, \tag{B.43}$$

where $\hat{K}(\theta) = K(\theta)/\int K(\theta)\sin\theta\mathrm{d}\theta$ is the spherically normalized latitudinal kernel. The dominant colatitude is the North-South region that gets sampled most by the kernel (e.g. the circles in Figure B.2.)

B.2.2. Albedo Variations

Figure 5.1 demonstrates that obliquity can influence a planet's apparent albedo on both rotational and orbital timescales. Quantifying these relations helps predict the obliquity constraints we may expect from real observations. We use a Monte Carlo approach, simulating planets with different maps and viewing geometries. We generate albedo maps from spherical harmonics, $Y_\ell^m(\theta,\phi)$, on the same 101×201 grid in colatitude and longitude from Section 5.2.2:

$$A(\theta,\phi) = \sum_{\ell=0}^{\ell_{\max}}\sum_{m=-\ell}^{\ell} C_\ell^m Y_\ell^m(\theta,\phi), \tag{B.44}$$

where $\ell_{\max}$ is chosen to be 3, each coefficient C_ℓ^m is randomly drawn from the standard normal distribution, and the composite map is scaled to the Earth-like range $[0.1, 0.8]$. Rotational and orbital changes in brightness are caused by East-West and North-South albedo markings, respectively, so we make three types of maps: East-West featured with $C_\ell^m(m \neq \ell) = 0$, North-South featured with $C_\ell^m(m \neq 0) = 0$, or no C_ℓ^m restrictions. For all maps with East-West features, we randomly offset the prime meridian. We generate 5,000 maps of each type.

For each map we randomly select an obliquity, solstice phase, inclination, and two orbital phases. Since inclination and orbital phase can be measured independent of photometry, we choose inclinations similar to planet Q, $i \in [50°,70°]$, and orbital phases $\{\xi_1, \xi_2\}$ with $\Delta\xi \in [110°, 130°]$. Both phases are also at least 30° from superior and inferior conjunction, which conservatively mimics an inner working angle at the selected inclinations. We assume the planet's rotational and orbital frequencies are known (Pallé et al., 2008; Oakley & Cash, 2009), and use the Earth-like ratio $\omega_{\rm rot}/\omega_{\rm orb} = 360$. We divide roughly one planet rotation centered on each orbital phase into 51 time steps, then define the normalized amplitude of rotational and orbital albedo variations, $\Lambda_{\rm rot}$ and $\Lambda_{\rm orb}$, as

$$\Lambda_{\rm rot} = \frac{A_{\xi_1}^{\rm high} - A_{\xi_1}^{\rm low}}{\bar{A}_{\xi_1}}, \tag{B.45}$$

$$\Lambda_{\rm orb} = |\bar{A}_{\xi_1} - \bar{A}_{\xi_2}| \left(\frac{\bar{A}_{\xi_1} + \bar{A}_{\xi_2}}{2}\right)^{-1}, \tag{B.46}$$

where $A_{\xi_1}^{\rm high}$ and $A_{\xi_1}^{\rm low}$ are the extreme apparent albedos around the first phase, and $\bar{A}$ is the mean apparent albedo of all time steps around a given phase. For each computed $\Lambda_{\rm rot}$ and $\Lambda_{\rm orb}$, we calculate the corresponding kernel width and absolute value change in dominant colatitude, from Appendix B.2.1. Figure B.3 shows the resulting distributions, where rotational and orbital information is colored red and blue, respectively. We find similar results when relaxing constraints on the inclination and orbital phases.

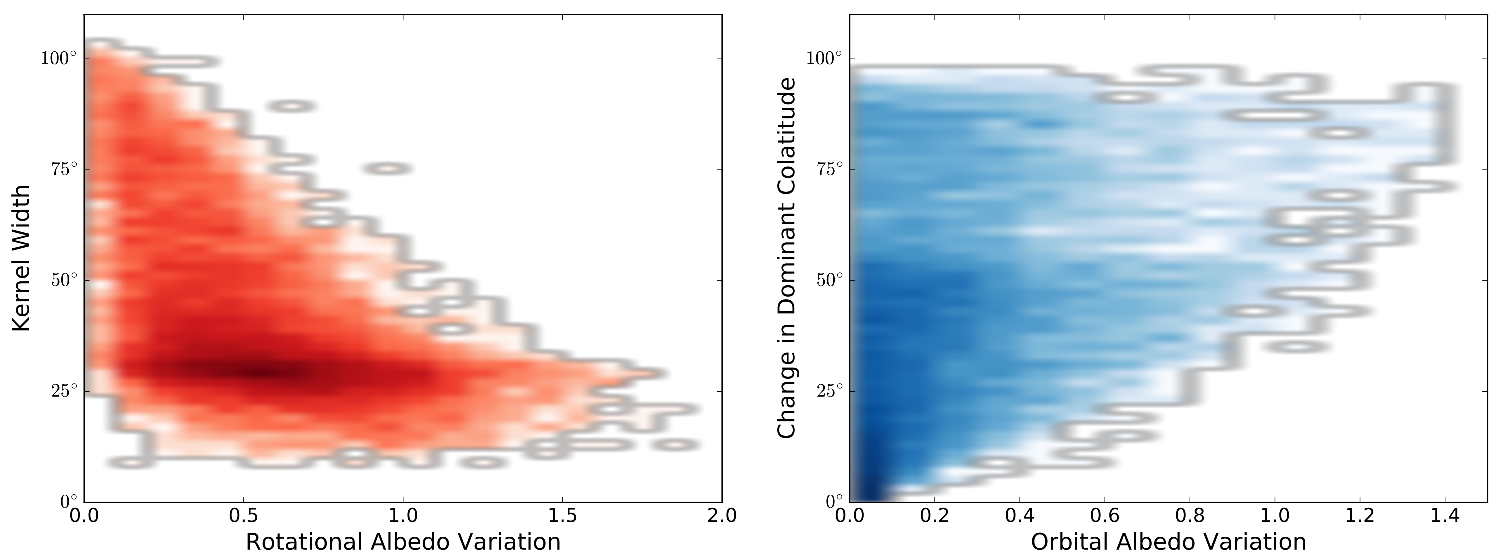


Figure B.3. Kernel-albedo distributions for two sets of 10,000 planets generated via Monte Carlo, comparing rotational properties in red and orbital properties in blue. Changes in dominant colatitude are absolute values, and planets are binned in $0.1 \times 2^\circ$ regions in both panels. Each color scale is logarithmic: the darkest red and blue bins contain 230 and 279 planets, respectively. The scatter in either kernel characteristic decreases as the corresponding albedo variation rises. We can estimate the uncertainty on a kernel width or change in dominant colatitude by analyzing these distributions.

We can estimate uncertainties on values of σ_ϕ and $|\Delta\bar{\theta}|$ using these distributions. The mean rotational and orbital variations are $\bar{\Lambda}_{\mathrm{rot}} \approx 0.54$ and $\bar{\Lambda}_{\mathrm{orb}} \approx 0.21$; the average kernel width and change in dominant colatitude are both roughly 38°. The full distributions have standard deviations of about 17° in σ_ϕ and 24° in $|\Delta\bar{\theta}|$, but roughly 5° and 7°, respectively, when considering only large variations. To predict constraints on obliquity obtained from real data, we will assume there are single- and dual-epoch observations of planet Q that have our mean variations $\bar{\Lambda}_{\mathrm{rot}}$ and $\bar{\Lambda}_{\mathrm{orb}}$. By considering samples only around these variations, we find about 10° and 20° standard deviations apiece in the kernel width and change in dominant colatitude. We use these standard deviations as uncertainties when creating the colored regions in Figure 5.6.

B.2.3. Peak Motion

Equations C1 and C2 from Cowan et al. (2009) describe the motion of the kernel peak, where specular reflection occurs, for any planetary system. These equations can be written for edge-on, zero-obliquity planets using Section B.1.3:

$$\cos\theta_{\mathrm{spec}} = \frac{1+\cos i}{\sqrt{2(1+\cos i)}}, \tag{B.47}$$

$$\begin{aligned} \tan\phi_{\mathrm{spec}} &= \frac{\sin(\xi - \omega_{\mathrm{rot}}t) + \sin(\phi_o(0) - \omega_{\mathrm{rot}}t)}{\cos(\xi - \omega_{\mathrm{rot}}t) + \cos(\phi_o(0) - \omega_{\mathrm{rot}}t)} \\ &= \frac{\sin(\omega_{\mathrm{orb}}t - \omega_{\mathrm{rot}}t) + \sin(\phi_o(0) - \omega_{\mathrm{rot}}t)}{\cos(\omega_{\mathrm{orb}}t - \omega_{\mathrm{rot}}t) + \cos(\phi_o(0) - \omega_{\mathrm{rot}}t)}. \end{aligned} \tag{B.48}$$

When finding ϕ_{spec} from Equation B.48, the two-argument arctangent must be used to ensure $\phi_{\mathrm{spec}} \in [-\pi, \pi)$. This also means it is difficult to simplify the equation with trigonometric identities.

Instead, we can explicitly write the argument of Equation B.48 in terms of the first meridian crossing, ξ_m, the earliest orbital phase after superior conjunction that the kernel peak recrosses the prime meridian:

$$\phi_{\text{spec}}(\xi;\xi_m) = \mp\frac{\pi}{2}\left(4\frac{\xi}{\xi_m} + \left[1 - \text{sgn}\left(\cos\frac{\xi}{2}\right)\right]\right), \tag{B.49}$$

where the leading upper sign applies to prograde rotation and vice versa. The first meridian crossing is related to the planet's frequency (or period) ratio by

$$\left|\frac{\omega_{\text{rot}}}{\omega_{\text{orb}}}\right| = \left|\frac{P_{\text{orb}}}{P_{\text{rot}}}\right| = \frac{1}{2}\left(\frac{4\pi}{\xi_m} \pm 1\right), \tag{B.50}$$

while the number of solar days per orbit is

$$N_{\text{solar}} = \left|\frac{\omega_{\text{rot}}}{\omega_{\text{orb}}}\right| \mp 1, \tag{B.51}$$

following the same sign convention. Note that the frequency/period ratios and the number of solar days do not have to be integers. We reiterate that Equations B.47–B.51 apply to *edge-on, zero-obliquity* planets.

Equation B.50 gives two frequency ratios for each first meridian crossing, one prograde and another retrograde that is smaller in magnitude by unity. Equation B.51 then states the corresponding difference in solar days is unity but reversed, making the longitudes of both kernel peaks in Equation B.49 analogous at each orbital phase. These two versions of the planet have East-West mirrored albedo maps and identical light curves: they are formally degenerate. An inclined, oblique planet has pro/retrograde versions that could be distinguished, as discussed in Section 5.4.5.